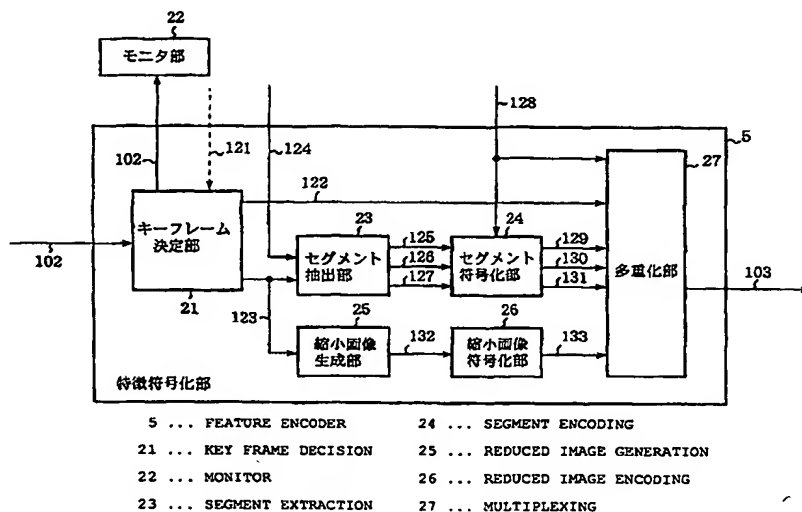




(51) 国際特許分類6 H04N 7/24, 5/93	A1	(11) 国際公開番号 WO00/45600 (43) 国際公開日 2000年8月3日(03.08.00)
(21) 国際出願番号 PCT/JP99/00403 (22) 国際出願日 1999年1月29日(29.01.99) (71) 出願人 (米国を除くすべての指定国について) 三菱電機株式会社 (MITSUBISHI DENKI KABUSHIKI KAISHA)[JP/JP] 〒100-8310 東京都千代田区丸の内二丁目2番3号 Tokyo, (JP) (72) 発明者 ; および (75) 発明者/出願人 (米国についてののみ) 関口俊一(SEKIGUCHI, Shunichi)[JP/JP] 山田悦久(YAMADA, Yoshihisa)[JP/JP] ジェームズ チャオ(JAMES, Chow)[US/JP] 浅井光太郎(ASAI, Kohtaro)[JP/JP] 〒100-8310 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内 Tokyo, (JP) (74) 代理人 弁理士 田澤博昭, 外(TAZAWA, Hiroaki et al.) 〒100-0013 東京都千代田区霞が関三丁目7番1号 大東ビル7階 Tokyo, (JP)		(81) 指定国 AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, HR, HU, ID, IL, IN, IS, JP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, 欧州特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), ARIPO特許 (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM) 添付公開書類 国際調査報告書

(54)Title: METHOD OF IMAGE FEATURE ENCODING AND METHOD OF IMAGE SEARCH**(54)発明の名称** 画像特徴符号化方法及び画像検索方法**(57) Abstract**

A feature encoding section (5) extracts a feature of a video signal (102), encodes it, and produces a feature stream (103). A feature identification section (11) collates a feature (109) encoded from the feature stream (103) with a search key (108) entered by a user to find video contents (111) that the user needs.

(57)要約・

特徴符号化部 5 は、ビデオ信号 1 0 2 の特徴量を抽出・符号化し、特徴ストリーム 1 0 3 を生成する。特徴同定部 1 1 は、特徴ストリーム 1 0 3 を復号した復号特徴量 1 0 9 とユーザからの検索キー 1 0 8 とのマッチングを行い、ユーザが求めたビデオコンテンツ 1 1 1 を検索する。

PCTに基づいて公開される国際出願のパンフレット第一頁に掲載されたPCT加盟国を同定するために使用されるコード(参考情報)

AE	アラブ首長国連邦	DM	ドミニカ	KZ	カザフスタン	RU	ロシア
AG	アンティグア・バーブーダ	DZ	アルジェリア	LC	セントルシア	SD	スーダン
AL	アルバニア	EE	エストニア	LI	リヒテンシュタイン	SE	スウェーデン
AM	アルメニア	ES	スペイン	LK	スリ・ランカ	SG	シンガポール
AT	オーストリア	FI	フィンランド	LR	リベリア	SI	スロヴェニア
AU	オーストラリア	FR	フランス	LS	レソト	SK	スロヴァキア
AZ	アゼルバイジャン	GA	ガボン	LT	リトアニア	SL	シエラ・レオネ
BA	ボスニア・ヘルツェゴビナ	GB	英国	LV	ラトヴィア	SN	セネガル
BB	バルバドス	GD	グレナダ	MC	モナコ	SZ	スワジランド
BE	ベルギー	GE	グルジア	MD	モルドヴァ	TD	チャード
BF	ブルキナ・ファソ	GH	ガーナ	MG	マダガスカル	TG	トーゴ
BG	ブルガリア	GN	ギニア	MK	マケドニア旧ユーゴスラヴィア	TJ	タジキスタン
BJ	ベナン	GR	ギリシャ		共和国	TM	トルクメニスタン
BR	ブラジル	GW	ギニア・ビサウ	ML	マリ	TR	トルコ
BY	ベラルーシ	HR	クロアチア	MN	モンゴル	TT	トリニダード・トバゴ
CA	カナダ	HU	ハンガリー	MR	モリタニア	TZ	タンザニア
CF	中央アフリカ	ID	インドネシア	MW	マラウイ	UA	ウクライナ
CG	コンゴ	IE	アイルランド	MX	メキシコ	UG	ウガンダ
CH	スイス	IL	イスラエル	MZ	モザンビーク	US	米国
CI	コートジボアール	IN	インド	NE	ニジェール	UZ	ウズベキスタン
CM	カメルーン	IS	アイスランド	NL	オランダ	VN	ヴェトナム
CN	中国	IT	イタリア	NZ	ニュージーランド	YU	ユーゴスラヴィア
CU	コスタ・リカ	JP	日本	PL	ポーランド	ZA	南アフリカ共和国
CY	キプロス	KE	ケニア	PT	ポルトガル	ZW	ジンバブエ
CZ	チェッコ	KG	キルギスタン	RO	ルーマニア		
DE	ドイツ	KP	北朝鮮				
DK	デンマーク	KR	韓国				

明 細 書

画像特徴符号化方法及び画像検索方法

技術分野

この発明は、アナログ又はデジタルで記録された映像、静止画などの画像データの特徴量を抽出・符号化し、符号化された特徴量を利用して画像データの検索を行う画像特徴符号化方法及び画像検索方法に関するものである。

背景技術

第1図は、電子情報通信学会論文誌D-I I, 1996年4月号(Vol. 79-D-I I, No. 4, pp. 476-483)に記載された従来の画像検索処理システムを説明するブロック図である。図において、91は、静止画像201を各セグメントに領域分割し、分割された各セグメントに対してキーワードを付与し、概念キーワード203とシーン記述キーワード204を出力する前処理部であり、92は、ユーザ97があらかじめ用意されているキーワード205を入力し、概念キーワード203とシーン記述キーワード204と照合して、静止画像201を検索する検索ツールである。

ここで、概念キーワード203は、各セグメントの色情報及び特徴量を示すキーワードであり、シーン記述キーワード204は、各セグメントの画像特徴量を、「位置」、「色」、「形状」、「大きさ」、「向き」等に関する述語を用いて表現したキーワードである。

第1図の前処理部91において、93は静止画像201を各セグメントに領域分割する領域分割部、94は、色情報に対してあらかじめ割り

当てられている概念キーワードを用いて、領域分割部 9 3 により分割された各セグメントの色及び特徴量から、概念キーワード 2 0 3 を抽出する概念キーワード抽出部、9 5 は、領域分割部 9 3 により分割された各セグメントの画像特徴量について、ユーザ 9 6 があらかじめ定義されている述語から選択して記述する述語記述 2 0 2 を入力することにより、シーン記述キーワード 2 0 4 を記述するシーン記述キーワード記述部である。

第 1 図の検索ツール 9 2 において、9 8 は、ユーザ 9 7 が選択したキーワード 2 0 5 と、前処理部 9 1 からの概念キーワード 2 0 3、シーン記述キーワード 2 0 4 から特徴量の同定を行う特徴同定部である。

次に動作について説明する。

静止画像 2 0 1 が前処理部 9 1 に入力されると、領域分割部 9 3 は静止画像 2 0 1 を各セグメントに領域分割する。概念キーワード抽出部 9 4 は、色情報に対してあらかじめ割り当てられている概念キーワードを用いて、分割された各セグメントの色及び特徴量から、概念キーワード 2 0 3 を抽出する。

また、シーン記述キーワード記述部 9 5 は、分割された各セグメントの画像特徴量について、ユーザ 9 6 からの述語記述 2 0 2 により、シーン記述キーワード 2 0 4 を記述する。

ユーザ 9 7 は、静止画像 2 0 1 を検索するにあたり、あらかじめ用意されている概念キーワード 2 0 3 やシーン記述キーワード 2 0 4 の中から選択したキーワード 2 0 5 を検索ツール 9 2 に入力する。特徴同定部 9 8 は、ユーザ 9 7 からのキーワード 2 0 5 と、前処理部 9 1 からの概念キーワード 2 0 3、シーン記述キーワード 2 0 4 から、特徴量の同定を行い、ユーザ 9 7 が求める静止画像 2 0 1 を検索する。

上記の画像検索処理システムは、静止画像 2 0 1 のみを対象としてい

るために、動画像を検索することが困難であるという課題があった。

また、各キーワードの記述方法・記憶方法などについて考慮されていないため、画像サーバとクライアント（検索ツール 9 2）とは、1 対 1 に対応している必要があり、ネットワークを介して、多くのユーザが様々な検索ツールを用いて、画像の検索を行う画像検索処理システムを構築することができないという課題があった。

この発明は上記のような課題を解決するためになされたもので、多くのユーザが、様々な検索ツールを用いて画像を検索することができる画像特徴符号化方法及び画像検索方法を得ることを目的とする。

発明の開示

この発明に係る画像特徴符号化方法は、画像フレームから所定の画像領域であるセグメントを抽出し、抽出した上記セグメントごとに上記セグメントを識別するセグメント番号を付与し、上記セグメントに代表色を割り当て、上記セグメントの上記画像フレームに対する面積率を算出し、上記セグメントごとに、上記代表色及び上記面積率を上記セグメントの特徴量として、上記セグメント番号と共に符号化して特徴ストリームを生成するものである。

このことによって、画像検索に使用する代表色及び面積率をセグメントの特徴量とした特徴ストリームを生成できるという効果を奏する。

この発明に係る画像特徴符号化方法は、画像フレームからセグメントを抽出する際に、色情報に基づく領域分類により上記セグメントを抽出し、上記セグメントに代表色を割り当てる際に、上記セグメントの抽出に用いた上記色情報を上記代表色として割り当てるものである。

このことによって、画像検索に使用する代表色及び面積率をセグメントの特徴量とした特徴ストリームを生成できるという効果を奏する。

この発明に係る画像特徴符号化方法は、画像フレームからセグメントを抽出する際に、隣接する画像フレーム間で上記セグメントの同定を行い、同じセグメントと判定されたセグメントに対して同一のセグメント番号を付与するものである。

このことによって、検索効率を向上可能な特徴ストリームを生成できるという効果を奏する。

この発明に係る画像特徴符号化方法は、画像フレームからセグメントを抽出する際に、隣接する画像フレーム間でセグメントの追跡を行い、同じセグメントと判定されたセグメントに対して動き情報を求め、上記セグメントごとに、上記動き情報を上記セグメントの特徴量として符号化して特徴ストリームを生成するものである。

このことによって、動き情報により画像フレームを検索可能な特徴ストリームを生成することができるという効果を奏する。

この発明に係る画像特徴符号化方法は、ビデオ信号のフレーム群から検索の手がかりとして適切な画像フレームをキーフレームとして抽出し、抽出したキーフレームからセグメントを抽出するものである。

このことによって、画像検索の手がかりになるキーフレームの特徴ストリームを生成することができるという効果を奏する。

この発明に係る画像特徴符号化方法は、キーフレームの所定領域ごとの画素値の平均値を求めて上記キーフレームの縮小画像を生成し、上記キーフレームごとに、上記縮小画像を上記キーフレームの特徴量として符号化して特徴ストリームを生成するものである。

このことによって、ラフにブラウズ可能な縮小画像を生成することができるという効果を奏する。

この発明に係る画像検索方法は、画像フレームを記憶する第1の記憶部と、上記第1の記憶部に記憶している画像フレームの特徴量を符号化

して特徴ストリームとして記憶する第２の記憶部を備えて画像検索を行うものにおいて、ユーザからの検索指示に基づき、上記第２の記憶部に記憶している上記特徴量を復号し、ユーザからの検索条件と上記復号した特徴量との同定を行い、上記第１の記憶部に記憶している画像フレームを検索するものである。

このことによって、ユーザは様々な検索ツールを用いて、画像フレームを検索することができるという効果を奏する。

この発明に係る画像検索方法は、第２の記憶部に記憶している特徴量として、画像フレームにおける所定の画像領域であるセグメントの代表色を含み、ユーザからの検索条件として上記代表色を含むものである。

このことによって、ユーザはセグメントの代表色により、画像フレームを検索することができるという効果を奏する。

この発明に係る画像検索方法は、第２の記憶部に記憶している特徴量として、画像フレームにおける所定の画像領域であるセグメントの上記画像フレームに対する面積率を含み、ユーザからの検索条件として上記面積率を含むものである。

このことによって、ユーザはセグメントの面積率により、画像フレームを検索することができるという効果を奏する。

この発明に係る画像検索方法は、第２の記憶部に記憶している特徴量として、隣接する画像フレームの動き情報を含み、ユーザからの検索条件として上記動き情報を含むものである。

このことによって、ユーザは動き情報により、画像フレームを検索することができるという効果を奏する。

この発明に係る画像検索方法は、第２の記憶部に記憶している特徴量として、画像フレームの縮小画像を含み、ユーザからの検索条件と上記復号した特徴量との同定を行い、上記縮小画像を上記ユーザに提示する

ものである。

このことによって、ユーザは縮小画像により、画像フレームを効率的に検索することができるという効果を奏する。

この発明に係る画像検索方法は、第2の記憶部に記憶している特徴量として、指定された被写体が画像フレームに存在するか否かの情報を含むものである。

このことによって、ユーザは画像データの検索効率を向上することができるという効果を奏する。

この発明に係る画像検索方法は、第2の記憶部に記憶している特徴量として、指定された被写体が以後の画像フレームに存在するか否かの情報を含むものである。

このことによって、ユーザは画像データの検索効率を向上することができるという効果を奏する。

この発明に係る画像検索方法は、第2の記憶部に記憶している特徴量として、指定された被写体が以前の画像フレームに存在するか否かの情報を含むものである。

このことによって、ユーザは画像データの検索効率を向上することができるという効果を奏する。

この発明に係る画像検索方法は、ユーザからの検索条件と復号した特徴量との同定を行う際に、優先度をユーザに提示するものである。

このことによって、ユーザは要求に即した画像データを迅速に検索できるという効果を奏する。

この発明に係る画像検索方法は、ユーザからの検索条件と復号した特徴量との同定を行う際に、ユーザからの複数の検索条件と復号した複数の特徴量との同定を行い、総合的に判断して、第1の記憶部に記憶している画像フレームを検索するものである。

このことによって、ユーザは画像データの検索効率を向上することができるという効果を奏する。

図面の簡単な説明

第 1 図は従来の画像検索処理システムを示すブロック図である。

第 2 図はこの発明が利用される典型的なシステム構成を示すブロック図である。

第 3 図はビデオコンテンツと特徴ストリームの対応を説明する図である。

第 4 図はこの発明の実施の形態 1 による特徴符号化部 5 の内部構成を示すブロック図である。

第 5 図はこの発明の実施の形態 1 による特徴符号化部 5 の動作を示すフローチャートである。

第 6 図は第 4 図におけるキーフレーム決定部 2 1 の内部構成を示すブロック図である。

第 7 図は第 4 図におけるキーフレーム決定部 2 1 の他の内部構成を示すブロック図である。

第 8 図はセグメントを説明する図である。

第 9 図はビデオシーケンスの中でのセグメントの位置づけを説明する図である。

第 10 図は第 4 図におけるセグメント抽出部 2 3 の内部構成を示すブロック図である。

第 11 図は M P E G - 4 のビデオオブジェクトプレーン (V O P) を説明する図である。

第 12 図は第 10 図のセグメント抽出部 2 3 を用いる場合の特徴符号化部 5 による特徴ストリーム 1 0 3 の構成を示す図である。

第 1 3 図はカラーマップテーブル 1 2 8 の具体例を示す図である。

第 1 4 図はこの発明の実施の形態 2 によるセグメント抽出部 2 3 の内部構成を示すブロック図である。

第 1 5 図は第 1 4 図のセグメント抽出部 2 3 を用いる場合の特徴符号化部 5 による特徴ストリーム 1 0 3 の構成を示す図である。

第 1 6 図はこの発明の実施の形態 3 によるセグメント抽出部 2 3 の内部構成を示すブロック図である。

第 1 7 図は第 1 6 図のセグメント抽出部 2 3 を用いる場合の特徴符号化部 5 による特徴ストリーム 1 0 3 の構成を示す図である。

第 1 8 図はこの発明の実施の形態 4 による特徴同定部 1 1 の内部構成を示すブロック図である。

第 1 9 図はこの発明の実施の形態 5 による優先度付き検索候補提示の例を示す図である。

発明を実施するための最良の形態

以下、この発明をより詳細に説明するために、この発明を実施するための最良の形態について、添付の図面に従って説明する。

実施の形態 1 .

この実施の形態 1 では、ビデオシーケンス中に含まれる被写体に着目し、各フレームに含まれる被写体領域を囲む矩形領域をセグメントと定義し、ビデオ信号の各フレーム中からセグメントを抽出して、その特徴量を求めて特徴ストリームを生成する処理を行う装置について説明する。

。

まず、前提として同装置を使用するシステムについて説明する。第 2 図は、この発明が利用される典型的なシステム構成を示すブロック図である。図において、1 はクライアント、2 はサーバ、3 はビデオコンテ

ンツ 1 1 1 を記憶しているビデオコンテンツ記憶部、4 は、ビデオコンテンツ記憶部 3 に記憶されているビデオコンテンツ 1 1 1 の中から、デジタル圧縮されたビットストリーム 1 0 1 を復号し、ビデオ信号 1 0 2 を出力する復号処理部である。

また、第 2 図において、5 は、復号処理部 4 により復号されたビデオ信号 1 0 2 を入力し、特徴ストリーム 1 0 3 を生成する特徴符号化部、6 は、ユーザ指示・設定 1 0 4 により、特徴符号化部 5 を制御するユーザ I / F、7 は特徴符号化部 5 により生成された特徴ストリーム 1 0 3 を記憶する特徴ストリーム記憶部である。ビデオコンテンツ記憶部 3 と特徴ストリーム記憶部 7 は、サーバ 2 内に別々に設置されているが、同一の記憶部に記憶されていても良い。

クライアント 1 において、8 は、ユーザ指示・設定 1 0 5 により、クライアント 1 の制御のために、検索制御信号 1 0 6 を出力するユーザ I / F、9 は、検索制御信号 1 0 6 を入力して、検索指示 1 0 7 と検索キー 1 0 8 を出力する検索処理部である。

さらに、第 2 図において、10 は、検索処理部 9 からの検索指示 1 0 7 により、特徴ストリーム記憶部 7 に記憶されている特徴ストリーム 1 0 3 を復号し、復号特徴量 1 0 9 を出力する特徴復号部、11 は、検索処理部 9 からの検索キー 1 0 8 と、特徴復号部 10 からの復号特徴量 1 0 9 を入力してマッチング処理を行い、特徴ストリーム番号を示すポインタ 1 1 0 を出力する特徴同定部である。111 は、特徴同定部 11 が出力したポインタ 1 1 0 により、ビデオコンテンツ記憶部 3 から出力されたビデオコンテンツである。

次に動作について説明する。

第 2 図に示されるシステムは、ビデオコンテンツ記憶部 3 に保持される各々のビデオコンテンツ 1 1 1 に対応づけられた特徴ストリーム 1 0

3に基づいて、ユーザが所望のビデオコンテンツ111を検索・ブラウズ（拾い読み）するシステムである。ここでは便宜上、サーバ・クライアントモデルを用い、検索システムをクライアント3、コンテンツ記憶・送出システムをサーバ4とみなして説明する。

クライアント3とサーバ4は、ネットワークを介して動作してもよいし、同一装置内で動作してもよい。第2図では、ネットワーク経由の動作は便宜上図示していないが、以下でネットワーク処理に必要な動作は適宜説明する。

（1）特徴ストリームの意義

第3図はビデオコンテンツ111と特徴ストリーム103の対応を説明する図である。図において、VCはビデオコンテンツ（Video Content）111を示し、FSは特徴ストリーム（Feature Stream）103を示している。VC及びFSそれぞれに、対応する番号を付しており、FS_nは、ビデオコンテンツVC_nの特徴量を記述した特徴ストリームである。VC_nとFS_nは、各ポインタ110により対応づけられている。

第3図に示すように、VC₁の第kフレームに「青い空」が含まれているとすれば、FS₁はVC₁の第kフレームに「代表色が青で、かつフレーム画面中の80%を占める領域」が存在することを記述する。このように、特徴量を属性として持つ領域を以下ではセグメントと呼ぶ。これにより、第2図におけるクライアント1から「代表色が青で、かつフレーム画面中の80%を占めるセグメント」という指定を行うことで、特徴ストリーム103から対応する特徴量を探し出し、対応するフレーム画面を含むビデオコンテンツ111を引き出すことを可能とする。

特徴量としては、代表色、サイズ、画面内位置、形状、位置の変化（動き）、明るさ・色の変化、コントラストなど様々な属性を定義するこ

とができる。もちろん、特徴量の指定方法は任意であり、クライアント 1 の検索処理部 9 に適当な解析機能を設けることで、ユーザは、より直感的な指定を行うようにすることもできる。

このように、ビデオコンテンツ (V C n) 1 1 1 中の特徴量を、特徴ストリーム (F S n) 1 0 3 に記述しておくことにより、ユーザは大量のビデオコンテンツ 1 1 1 のデータの中から、直感的な検索キー 1 0 8 を用いて、所望のコンテンツ 1 1 1 を探し出すことが容易になる。このような検索方法は、極めて膨大な数のビデオコンテンツ 1 1 1 を含むビデオライブラリ・データベース、ビデオテープなどから、所望のコンテンツ 1 1 1 を探し出す処理を効率化することができる。

(2) 特徴ストリームの生成

特徴ストリーム 1 0 3 の生成処理は、ビデオコンテンツ (V C n) 1 1 1 に対応する特徴ストリーム (F S n) 1 0 3 を生成し、特徴ストリーム記憶部 7 に格納する処理に相当する。この処理を行うのは、復号処理部 4、特徴符号化部 5、ユーザ I / F 6 である。復号処理部 4 は、ビデオコンテンツ (V C n) 1 1 1 がデジタル圧縮されたビットストリーム 1 0 1 の形式で記憶されている場合にのみ必要で、ビデオ信号 1 0 2 を出力する。ビデオコンテンツ (V C n) 1 1 1 が直接表示可能な画像データである場合には、復号処理部 4 は不要である。

特徴符号化部 5 は、ビデオ信号 1 0 2 に基づいて、特徴ストリーム (F S n) 1 0 3 を生成し、特徴ストリーム記憶部 7 に格納する。特徴ストリーム 1 0 3 の生成処理の詳細は後述する。

(3) 検索処理

検索処理はクライアント 1 によって行われる。クライアント 1 は、ユーザが、ビデオコンテンツ記憶部 3 に記憶されているビデオコンテンツ 1 1 1 の中から、所望のコンテンツ 1 1 1 を検索するための処理部であ

る。ユーザは、ユーザ指示・設定 105 により、クライアント 1 のユーザ I / F 8 を介して、特徴ストリーム 103 として符号化されている特徴量に対応する値を求めるために、検索制御信号 106 を検索処理部 9 に与える。

第 3 図を例にとり、「青い空」が含まれるビデオコンテンツ (VC 1) 111 を、すべてのビデオコンテンツ (VC n) 111 の中から検索する例を簡単に説明する。この検索処理は、すべての特徴ストリーム (FS n) 103 に対して、「青い空」を示す特徴量をサーチし、それを含む特徴ストリーム (FS n) 103 に対応するビデオコンテンツ (VC n) 111 を特定する処理である。検索処理部 9 は、検索制御信号 106 から検索すべき特徴量の値を求めて、検索キー 108 として送出すると同時に、検索指示 107 をアクティブにする。

もし、クライアント 1 とサーバ 2 がネットワーク接続されているとすれば、検索指示 107 と検索キー 108 は、ネットワークを通じて、サーバ 2 及び特徴量同定処理を行う装置 (特徴復号部 10 と特徴同定部 11) に伝送される。

検索指示 107 がアクティブになると、特徴復号部 10 は特徴ストリーム記憶部 7 から順次特徴ストリーム (FS n) 103 を取り出し、特徴ストリーム 103 から特徴量の値を復号する。その結果、得られた復号特徴量 109 は、特徴同定部 11 において、検索キー 108 とマッチング処理される。ここで、検索キー 108 と一致する復号特徴量 109 が見つかった場合は、対応する特徴量を含む特徴ストリーム 103 のポインタ 110 から、ビデオコンテンツ (VC n) 111 を特定する。第 3 図の例では、「青い空」を含むのが VC 1 のみなので、FS 1 における第 k フレーム目のセグメントの中に、検索キー 108 に一致する復号特徴量 109 が見つかり、VC 1 の第 k フレームが検索結果として出力

される。

ここで、特徴量同定処理を行う装置（特徴復号部 10 と特徴同定部 11）は、クライアント 1 に含まれていてもよいし、サーバ 2 に含まれていてもよいし、ネットワーク中に存在する別の装置に含まれていてもよい。もちろん、クライアント 1 とサーバ 2 とが同一装置になっている場合は、同一装置内に含まれる。

検索結果として出力されたビデオコンテンツ 111 は、クライアント 1 に送られ、ユーザはユーザ I / F 8 により、その画像の内容をブラウズできる。また、「青い空」を示す特徴量によって、検索結果として複数のビデオコンテンツ 111 が見つかった場合でも、ユーザ I / F 8 に表示することで、複数のビデオコンテンツ 111 をブラウズするようにすることもできる。このようなシステムによって、ユーザは直接ビデオコンテンツ 111 の中身をすべてブラウズすることなく、所望のセグメントを含むビデオコンテンツ 111 だけに絞り込んで内容を確認することができ、検索の効率が向上する。

（４）インタラクティブ機能

同システムにおいては、特徴符号化部 5 がビデオ信号 102 から特徴ストリーム 103 を生成する際に、ユーザ指示・設定 104 により、ユーザ I / F 6 を介して、ビデオコンテンツ（VCn）111 に含まれるセグメントを定義するための各種の条件設定や、セグメントの追加・削除、あるいはセグメントを定義するフレーム位置設定など、ユーザが任意に特徴ストリーム 103 の生成処理を制御できる。

（５）特徴ストリームの伝送・配信

特徴ストリーム（FSn）103 は、対応するビデオコンテンツ（VCn）111 へのポインタ 112 を持つ限り、ビデオコンテンツ 111 と共に、サーバ 2 に格納しておく必要はなく、どこに存在していてもよ

い。例えば、特徴ストリーム 103 だけが含まれる CD-ROM があり、クライアント 1 において、当該 CD-ROM を読むことにより、特徴ストリーム 103 に対応するビデオコンテンツ 111 の場所を特定することもできる。この場合、例えば特徴ストリーム 103 がビデオコンテンツの URL (Uniform Resource Locator) などを保持していればよい。

また、一般に、特徴ストリーム 103 は、ビデオコンテンツ 111 のデータ量に比べて小さいので、より小さな蓄積メディアに格納して、ノート型 PC (Personal Computer) や PDA (Personal Digital Assistants) などの携帯端末で利用するということも可能である。

さらに、特徴ストリーム 103 は、ビデオコンテンツ 111 に付加して、ネットワークを介して伝送・配信することもできる。上記 (4) で述べたインタラクティブ機能により、特徴ストリーム 103 を受信した側で、再度特徴ストリーム 103 の内容を加工・編集して、再利用・再配信もできるので、あらゆるメディア間で検索の自由度を失うことなく、ビデオコンテンツ 111 を自由に流通することができる。

次に、特徴ストリーム 103 の生成処理について詳しく説明する。

特徴ストリーム 103 の生成処理は、上述したように、主に特徴符号化部 5 において行われる。第 4 図は実施の形態 1 による特徴符号化部 5 の内部構成を示すブロック図である。図において、21 は、ビデオ信号 102 を入力して、キーフレーム設定指示 121 に基づきキーフレームを決定し、キーフレーム番号 122 とキーフレーム画像 123 を出力するキーフレーム決定部、22 は、キーフレーム決定部 21 からのビデオ信号 102 をモニタするモニタ部である。

また、第 4 図において、23 は、キーフレーム決定部 21 が決定した

キーフレーム画像 1 2 3 からセグメントを抽出し、セグメントの代表色 1 2 5, セグメント番号 1 2 6 を出力すると共に、例えばビデオコンテンツ 1 1 1 の復号時に出力されるキーフレームサイズ 1 2 4 により、セグメントのサイズ情報 1 2 7 を出力するセグメント抽出部である。

さらに、第 4 図において、2 4 は、セグメント抽出部 2 3 からのセグメントの代表色 1 2 5 を、カラーマップテーブル 1 2 8 を用いて符号化し、符号化されたセグメントの代表色 1 2 9 を出力すると共に、セグメント番号 1 2 6, セグメントのサイズ情報 1 2 7 を符号化し、符号化されたセグメント番号 1 3 0, 符号化されたセグメントのサイズ情報 1 3 1 を出力するセグメント符号化部である。

さらに、第 4 図において、2 5 は、キーフレーム画像 1 2 3 より、縮小画像を生成し、縮小画像データ 1 3 2 を出力する縮小画像生成部、2 6 は、縮小画像データ 1 3 2 を符号化し、符号化された縮小画像データ 1 3 3 を出力する縮小画像符号化部である。

さらに、第 4 図において、2 7 は、カラーマップテーブル 1 2 8, 符号化されたセグメントの代表色 1 2 9, 符号化されたセグメント番号 1 3 0, 符号化されたセグメントのサイズ情報 1 3 1, 符号化された縮小画像データ 1 3 3 を多重化する多重化部である。

次に特徴符号化部 5 の動作について説明する。

第 5 図は実施の形態 1 による特徴符号化部 5 の動作を示すフローチャートである。

(A) キーフレームの決定

まず、キーフレーム決定部 2 1 において、ビデオコンテンツ 1 1 1 中のキーとなるフレーム画像であるキーフレームを決定する（ステップ S T 1）。キーフレームは、ビデオコンテンツ 1 1 1 中の意味内容の変化点（シーンチェンジ）となるフレームや、ユーザが特に検索時の特徴点

としたいフレームと定義する。

第6図は第4図におけるキーフレーム決定部21の内部構成を示すブロック図である。第6図において、31は、ビデオ信号102のフレーム番号をカウントするフレームカウンタ、32は、キーフレーム設定指示121により、キーフレーム番号122を出力するスイッチ、33は、キーフレーム設定指示121により、キーフレーム画像123を出力するスイッチである。第6図では、ユーザがビデオ信号102をモニタ部22によりモニタリングしながら、所望のフレームをキーフレームとして指定する。

ビデオ信号102は、1フレームごとにキーフレーム決定部21に入力され、フレームカウンタ31により、フレーム番号がカウントされる。フレームカウンタ31は、ビデオ信号102の最初の入力時点でゼロにリセットされる。

また、ビデオ信号102は同時にモニタ部22に送られ、ユーザはモニタ部22に表示されるビデオ信号102を見ながら、キーフレームとすべきフレームを選択する。選択指示は、キーフレーム設定指示121をアクティブにすることで行う。キーフレーム設定指示121がアクティブになると、スイッチ32は選択したキーフレームのフレーム番号をキーフレーム番号122として出力し、スイッチ33は選択したキーフレームのキーフレーム画像123を出力する。

第7図は第4図におけるキーフレーム決定部21の他の内部構成を示すブロック図である。第7図において、41はビデオ信号102からシーンチェンジを検出し、キーフレーム画像123とキーフレーム設定指示141を出力するシーンチェンジ検出部、42は、キーフレーム設定指示141により、キーフレーム番号122を出力するスイッチである。フレームカウンタ31は、第6図におけるフレームカウンタ31と同

一である。ここでは、キーフレームは、シーンチェンジ点に位置するフレームとしており、キーフレーム決定部 2 1 は、シーンチェンジ検出部 4 1 を備えている。

ビデオ信号 1 0 2 は、第 6 図に示すキーフレーム決定部 2 1 と同様に、フレームカウンタ 3 1 に入力されて、フレーム番号がカウントされると共に、シーンチェンジ検出部 4 1 に入力されて、シーンチェンジ検出が実施される。シーンチェンジ検出の方法については、長坂、宮武、上田「カットの時系列コーディングに基づく映像シーンの実時間識別法」(電子情報通信学会論文誌、D-I I, V o l. J 7 9-D-I I, N o. 4, p p 5 3 1-5 3 7) などの先行技術があるが、この発明では具体的なアルゴリズムは対象外とする。

シーンチェンジが検出されると、キーフレーム設定指示 1 4 1 がアクティブになり、スイッチ 4 2 は、これに従って、その時点でのフレームカウント値をキーフレーム番号 1 2 2 として出力する。また、シーンチェンジ検出部 4 1 は、検出されたシーンチェンジフレームを、キーフレーム画像 1 2 3 として出力する。

また、図示はしていないが、一定間隔で現れるイントラ符号化フレームを、キーフレームに設定することも可能である。例えば、M P E G (M o v i n g P i c t u r e E x p r e t s G r o u p) - 1 や M P E G - 2、M P E G - 4 では、フレーム全体を、フレーム間予測を使わずに符号化するイントラ符号化フレームモードがあり、特にイントラ符号化フレームは、ビデオコンテンツ 1 1 1 の符号化中に周期的に挿入して、ランダムアクセスのポイントなどに利用することが多い。従って、キーフレームとしての要件を満たしていると考えられる。

(B) セグメントの抽出

キーフレーム画像 1 2 3 が選択されると、第 4 図におけるセグメント

抽出部23が、キーフレーム画像123の中に含まれるセグメントを抽出する。ここで、セグメントとは「キーフレーム上の指定領域、又は、ビデオコンテンツ111中に現れる被写体などの物体（＝オブジェクト）に対応したキーフレーム中の領域」と定義する。

第8図はセグメントを説明する図である。第8図で、赤い車がビデオコンテンツ111の複数のフレームにわたって現れる場合、赤い車がオブジェクトに相当し、セグメントはキーフレーム画像123上に現れるオブジェクト領域（車の外接四角形であらわす）である。ここでは、セグメントは、その領域の代表色とサイズ情報（フレーム全体に対するセグメント領域の面積率）という2つの特徴量を持っている。

セグメントの代表色が特徴ストリーム103の中に記述されていると、ユーザは「赤いセグメント」、「青いセグメント」を含むビデオコンテンツ111といった形式で、特徴ストリーム103の中の値と検索したい値とのマッチングをとることにより、ある程度、自動的に要求に合致するキーフレームを探し出すことができる。

また、サイズ情報については、セグメントがキーフレームの中で、どれくらいの面積を占めているかを記述することができるので、キーフレームにおけるセグメントの重要度の一側面を表現することができる。例えば、「画面いっぱいのサイズで、肌色の代表色を持つセグメント」などの指定の仕方により、画面いっぱいに映る顔画像を含むキーフレームを検索するといったことが、ほぼ自動的に行える。また、サイズ情報はサイズ計測の起点となる矩形の左上隅頂点や矩形の重心位置などの位置情報を含んでも良い。

第9図はビデオシーケンスにおけるセグメントの位置づけを説明する図である。第9図では、オブジェクトである「赤い車」が時間の経過に従って、画面の左から徐々にズームアップしながら、画面の右方向へ移

動するシーンを示しており、キーフレーム毎にセグメントの特徴量に変化する様子を示している。第9図では、第8図に示した各キーフレームの特徴量のほか、オブジェクトの動きとして、セグメントにおける左上隅の頂点（セグメント位置）の移動量を、時間方向の特徴量として持つことを示している。これによって、「画面の左から右へ移動するオブジェクト」という検索の仕方が可能になる。

また、例えばMPEG-4ビデオ符号化方式（ISO/IEC, JTC1/SC29/WG11, N2202）で規定されているビデオオブジェクトも、本定義におけるオブジェクトとみなすことができる。この場合、セグメントはMPEG-4ビデオのビデオオブジェクトプレーン（VOP）に対応する。この場合、厳密な定義は異なるが、VOPの水平・垂直サイズは、セグメントの矩形の水平・垂直サイズに、概念的には対応している。それに対し、MPEG-1やMPEG-2では、もともとオブジェクトの概念がないため、キーフレームからの抽出を行うことによって始めてセグメントが決定される。

セグメント抽出処理は、キーフレーム画像123の中から、以上のようなセグメントを切り出し、その特徴量を決定・獲得する処理であり、第4図のセグメント抽出部23で行われる。第10図は第4図におけるセグメント抽出部23の内部構成を示すブロック図である。第10図において、51は、キーフレーム画像123を入力し、セグメント番号126、セグメント内画像サンプル値151、矩形で示したセグメントの水平・垂直サイズ152（画素数×ライン数）を出力するセグメント抽出処理部、52は、セグメント内画像サンプル値151により、代表色125を割り当てる代表色割当部、53は、セグメントの水平・垂直サイズ152と、キーフレームサイズ124により、面積率で示されるサイズ127を算出するサイズ算出部である。

第 1.0 図におけるセグメント抽出部 2 3 は、まず、セグメント抽出処理部 5 1 において、キーフレーム画像 1 2 3 から、既存の領域分割処理などの方法によりセグメントを切り出す（第 5 図のステップ S T 2）。この例では、セグメントとオブジェクトの関係を厳密に結びつけないものとしている。すなわち、この例では、セグメントは単なるキーフレーム中の特定画像領域であり、キーフレーム間でオブジェクトとして関連付けるための機構を備えず、個々のキーフレームで独立したセグメントを定義する。

セグメント切り出し方法としては、例えば、色成分空間で類似色を集めるクラスタリングを行うなどの方法があるが、ここでは、その具体的な実現手法については対象外とし、セグメント抽出処理部 5 1 により、特定の意味内容を持つ画像領域（セグメント）が、外接四角形に囲まれた形で得られるものとする。

セグメント抽出処理部 5 1 は、抽出した画像領域（セグメント）ごとに順次番号をつけてカウントし、カウント値をセグメント番号 1 2 6 として出力する（ステップ S T 3）。

また、セグメント抽出処理部 5 1 は、セグメント内画像サンプル値 1 5 1 を代表色割当部 5 2 に出力し、代表色割当部 5 2 は、代表色 1 2 5 を求める（ステップ S T 4）。例えば、セグメント内画像サンプル値 1 5 1 が、各 8 ビットの R G B 表現になっている場合、セグメント中の R, G, B の各空間での平均値をとり、得られる R G B の平均値のセットを代表色として割り当てるなどの処理を行う。または、セグメント中の本質的な領域に含まれる画素をあらかじめ特定し、その内部での平均をとるなどの方法を用いることもできる。

M P E G - 4 の V O P をセグメントとするならば、セグメント中の本質的な領域は、V O P の形状を表現するアルファプレーンによって得ら

れる。第 11 図は M P E G - 4 のビデオオブジェクトプレーン (V O P) を説明する図であり、 V O P の画素データとアルファプレーンとの関係を示している。図に示すように、アルファプレーンとは、 V O P を囲む外接四角形のうち、オブジェクト内画素の位置の時に 2 5 5 を割り当て、それ以外の時に 0 を割り当てた 2 値のプレーンである。0 と 2 5 5 の変化点が V O P の形状をあらわす。つまり、アルファプレーンが 2 5 5 であるような位置の画素だけを用いて代表色を求める、という処理が可能となる。

このほか、代表色の求めかたとして、セグメント抽出処理部 5 1 で色情報に基づいた切り出しを行い、クラスタリングの結果として、セグメントに割り当てられた色情報を用いる方法もある。

また、第 10 図において、セグメント抽出処理部 5 1 は、矩形で示したセグメントの水平・垂直サイズ 1 5 2 (画素数×ライン数) を、サイズ算出部 5 3 に出力する。サイズ算出部 5 3 は、別途入力されるキーフレームの水平・垂直サイズ 1 2 4 を用いて、セグメントの当該キーフレーム上に占める面積率を計算し、サイズ 1 2 7 として出力する (ステップ S T 5)。例えば、セグメントの水平・垂直サイズ 1 5 2 が (1 7 6 画素×1 4 4 ライン) で、キーフレームのサイズ 1 2 4 が (3 5 2 画素×2 8 8 ライン) である場合、サイズ算出部 5 3 により算出された面積率 (サイズ 1 2 7) は 2 5 % となる。

(C) セグメントの符号化

第 4 図において、セグメント抽出部 2 3 で得られた特徴量 (代表色 1 2 5, セグメント番号 1 2 6, サイズ 1 2 7 など) は、セグメント符号化部 2 4 に入力され、多重化部 2 7 により特徴ストリーム 1 0 3 の形式に符号化される (ステップ S T 6)。

第 12 図は第 10 図のセグメント抽出部 2 3 を用いる場合の特徴符号

化部 5.による特徴ストリーム 1 0 3 の構成を示す図である。ここで特徴ストリーム 1 0 3 は、ビデオコンテンツ 1 1 1, キーフレーム、セグメントの階層構造を有する。ビデオコンテンツ 1 1 1 の階層では、フレーム番号ビット長やカラーマップテーブル 1 2 8 などのビデオコンテンツ自体に固有の情報をヘッダとして持ち、それに続いて K 枚のキーフレーム階層のデータを含む。ここで、各キーフレームの時間位置（キーフレーム番号 1 2 2）を各キーフレームのヘッダに多重化するため、そのビット長を規定するために最大フレームカウント値が表現可能なビット長を、ビデオコンテンツ 1 1 1 のヘッダに含めるものとする。さらに、各セグメントにおける代表色の符号化のため、ビデオコンテンツ 1 1 1 を通じて用いられる色階調を規定するカラーマップテーブル 1 2 8 を含める。

k 番目のキーフレームを $KF(k)$ と表記すると、 $KF(k)$ はビデオコンテンツ 1 1 1 の中での時間位置（キーフレーム番号 1 2 2）と、その画面上に存在するセグメントの数（M）をヘッダとして持ち、それに続いて M 個のセグメント階層のデータを含む。また、ブラウジングのために後述の縮小画像のデータを持っている。m 番目のセグメントを $SG(m)$ と表記すると、 $SG(m)$ は対応する代表色 1 2 5 とサイズ 1 2 7 から構成される。代表色 1 2 5 として、カラーマップテーブル 1 2 8 のインデックス（index）値を符号化する。

第 1 3 図はカラーマップテーブル 1 2 8 の具体例を示す図である。これは、X-Windows システムで用いられるカラーマップテーブル 1 2 8 を例にとったものであり、他にも様々な表現方法が考えられるが、ここでは、具体的なカラーマップテーブル 1 2 8 のデータ形式の定義については対象外とする。

このように、一つのインデックスに R, G, B の値のセットが対応し

ており、インデックスの数（ n ）を増やすことによって、階調を増やすことができる。

サイズ 1 2 7 は面積率であり、高々 1 ～ 1 0 0（％）の範囲の値であるため、7 ビットあれば表現可能である。

第 4 図において、セグメント符号化部 2 4 により、最終的に得られた各特徴量符号化データは、多重化部 2 7 に送られ、第 1 2 図で示される特徴ストリーム 1 0 3 に追加される。セグメント抽出部 2 3 によるセグメントの特徴量算出、及びセグメント符号化部 2 4 による符号化は、キーフレーム内で切り出されたセグメント数の分だけ繰り返す（第 5 図のステップ S T 7）。

（D）縮小画像の生成

第 4 図において、縮小画像生成部 2 5、縮小画像符号化部 2 6 は、キーフレームの概要画像（以下、サムネイル又はサムネイル画像と呼ぶ）を、キーフレームの特徴量として特徴ストリーム 1 0 3 に添付する目的で備えられる。キーフレームにアイコン的なサムネイル画像が用意されていれば、ユーザはデジタル圧縮されたコンテンツを直接復号して再生する必要がなくなるので、ブラウズの効率を向上できる。

ここでは、サムネイル画像として、キーフレーム画像 1 2 3 の $N \times N$ 画素ごとの平均値（D C 値）を求めることにより、縮小画像を生成し（第 5 図のステップ S T 8）、その値を適当な符号化方式により符号化する（ステップ S T 9）。

第 4 図において、縮小画像生成部 2 5 は、D C 値の縮小画像データ 1 3 2 を求める処理を行い、縮小画像符号化部 2 6 は、得られた D C 値の縮小画像データ 1 3 2 を圧縮符号化し、符号化された縮小画像データ 1 3 3 を出力する。

キーフレーム画像 1 2 3 は、すでに非可逆圧縮符号化されたビットス

ストリームからの復号画像である場合が多いので、縮小画像符号化部 26 での圧縮符号化は、DPCM (Differential Pulse Code Modulation) などの簡易で圧縮率の低い符号化を用いることが適当である。N×N画素ごとのDC値を求めることで、少なくともサンプル数は $1/N^2$ まで削減でき、特徴ストリーム 103 に追加しても、大きな符号量的負荷にならないようにすることができる。

符号化された縮小画像 133 は多重化部 27 に送られ、第 12 図の形式で特徴ストリーム 103 に多重化される。

以上のように、この実施の形態 1 によれば、上記特徴符号化部 5 の構成を取ることににより、ユーザはビデオコンテンツ 111 の特徴を記述した特徴ストリーム 103 を生成することができる。また、ビデオコンテンツ 111 から、マニュアル的又は自動的にキーフレームを設定することができ、さらに各キーフレーム中に存在する特定の画像領域（セグメント）に代表色 125 とそのサイズ 127 などの特徴量として設定することができる。これらの特徴量を検索キーとすることで、ビデオコンテンツの検索処理を、ある程度自動化することができ、自動検索によって絞り込んだ候補を、さらにサムネイル画像でラフにブラウズしたりできるので、ビデオコンテンツの検索効率を向上することができる。

さらに、この実施の形態 1 におけるセグメントの定義は、フレーム画像を静止画とみなして考えることができるので、膨大な静止画像ライブラリから、所望の画像を検索する場合にも利用可能である。静止画の場合、第 12 図のキーフレームの階層が最上位階層になる。

実施の形態 2.

第 14 図はこの発明の実施の形態 2 によるセグメント抽出部 23 の内部構成を示すブロック図である。図において、61 は、キーフレーム画

像 1 2 3 よりセグメントデータ 1 6 1 を抽出するセグメント抽出処理部、6 2 は、参照画像メモリ 6 3 に記憶されている過去のキーフレーム画像 1 2 3 とマッチング処理を行いセグメントを同定するセグメント同定処理部である。

この例では、セグメントはフレーム間でマッチングをとり、ビデオコンテンツ 1 1 1 内のオブジェクトと対応づけられる。すなわち、セグメントはキーフレームで閉じることなく、ビデオコンテンツ 1 1 1 における被写体のある瞬間を表す画像領域として抽出される。

セグメント抽出処理部 6 1 により、キーフレーム画像 1 2 3 から複数のセグメントデータ 1 6 1 が抽出されると、セグメント同定処理部 6 2 において、各セグメントについて、参照画像メモリ 6 3 における過去のキーフレーム画像 1 2 3 との間でマッチング処理を行い、セグメントの同定を行う。この際に同定されたセグメントについては、過去のセグメントと同一番号を付してセグメント番号 1 2 6 として出力する。また、過去のセグメントと一致しなかった場合は、新規セグメントとして新しいセグメント番号 1 2 6 を与えて出力する。

また、セグメント同定処理部 6 2 は、セグメント内画像サンプル値 1 5 1、セグメントの水平・垂直サイズ 1 5 2 を出力し、代表色割当部 5 2、サイズ算出部 5 3 は、第 1 0 図と同様に、それぞれ代表色 1 2 5、サイズ 1 2 7 の算出を行う。

第 1 5 図は第 1 4 図のセグメント抽出部 2 3 を用いる場合の特徴符号化部 5 による特徴ストリーム 1 0 3 の構成を示す説明図である。ここでは、セグメントがビデオコンテンツ 1 1 1 内のオブジェクトに対応づけられるため、セグメント数 (M) は、ビデオコンテンツ 1 1 1 の階層のヘッダに置かれる。

セグメント S G (m) は F l a g (1) を追加で持つ。F l a g (1

)は、SG (m) がKF (k) 上に存在するか否かを示す。各キーフレームは常にM個のセグメントを符号化するものとし、実際にSG (m) がKF (k) に存在しない場合、Flag (1) をOFFにして、代表色125, サイズ127を符号化しないようにする。また、このFlag (1) は、第4図における多重化部27により付加される。

また、SG (m) がKF (k) には存在するがKF (k-1) には存在しない、つまりフレームkにおいて登場する、という場合に特定の登場フラグで示すことや、SG (m) がKF (k) には存在するがKF (k+1) には存在しない、つまりフレームkを最後に退場する、という場合に特定の退場フラグで示すということも、フラグの追加により可能である。

最終的に得られた各特徴量符号化データは、多重化部27に送られ、第15図の特徴ストリーム103の形式で出力される。

以上のように、この実施の形態2によれば、第14図のセグメント抽出部23の構成を取ることににより、ビデオコンテンツ111内のオブジェクトと複数のキーフレームのセグメントを関連付けて、セグメントの数がオブジェクトの個数に対応して出力されるので、ユーザが求めるビデオコンテンツ111を効率的に検索することができる。

実施の形態3.

第16図は実施の形態3によるセグメント抽出部23の内部構成を示すブロック図である。図において、71は、キーフレーム画像123を入力し、参照画像メモリ63を参照して、セグメント番号126, セグメント内画像サンプル値151, セグメントの水平・垂直サイズ152を出力すると共に、動き情報171を出力するセグメント追跡処理部である。代表色割当部52, サイズ算出部53は第14図と同一である。

この例では、セグメントはビデオコンテンツ 1 1 1 を通じて、オブジェクトをトラッキング（追跡）した結果、各キーフレーム上に現れる画像領域として得られる。オブジェクトトラッキングは、セグメント追跡処理部 7 1 において行われる。

オブジェクトトラッキングの方法については、様々な手法が提案されており、ここでは、その方法自体は対象外とする。適当なアルゴリズムを用いることにより、オブジェクトが途中で画面から消えたり再出現したりといったレベルまで追跡できるものと仮定する。

結果として得られる各キーフレーム画像上のセグメントについて、第 1 0 図や第 1 4 図に示すセグメント抽出部 2 3 と同様に、代表色 1 2 5，サイズ 1 2 7 が算出されるほか、同一オブジェクトである限りは、同一のセグメント番号 1 2 6 が出力される。また、ここでは、トラッキングの結果として、セグメントの移動量が得られるものとし、それを動き情報 1 7 1 として出力する。

第 4 図の構成では、セグメント抽出部 2 3 から、動き情報 1 7 1 は出力されていないが、第 1 6 図のセグメント抽出部 2 3 を用いる場合は、動き情報 1 7 1 も出力するものと考え、これがセグメント符号化部 2 4 で符号化されるものとする。

第 1 7 図は第 1 6 図のセグメント抽出部 2 3 を用いる場合の特徴符号化部 5 による特徴ストリーム 1 0 3 の構成を示す説明図である。第 1 5 図と異なるのは、セグメントの構造だけで、この場合は、初めて現れたセグメントかどうかを示すフラグ F l a g (2) を設け、F l a g (2) がアクティブの場合に、動き情報 1 7 1 をまとめて多重化する構成になっている。

このような構造をとることにより、ビデオコンテンツ 1 1 1 の中でオブジェクトが新規発生した個所に、それ以降の動きの履歴がすべて含ま

れているため、たとえば「左から右へ移動した」などの検索キーに、すばやく対応することが可能となる。第 17 図に図示はしていないが、もちろん、動き情報 171 は各キーフレーム内のセグメントごとに付加してもよい。

第 4 図において、セグメント符号化部 24 により最終的に得られた各特徴量符号化データは、多重化部 27 に送られ、第 17 図の特徴ストリーム 103 の形式として出力される。

以上のように、この実施の形態 3 によれば、動き情報 171 を設定することにより、各キーフレーム間で移動するオブジェクトを検索することができる。

実施の形態 4.

この実施の形態では、第 2 図のクライアント 1 によるビデオコンテンツ 111 の検索処理について説明する。ユーザは、クライアント 1 のユーザ I/F 8 を用いて、ユーザ指示・設定 105 により、検索したいビデオコンテンツ 111 の大まかなパラメータの入力操作を行う。入力方法としては、クライアント 1 が、あらかじめ用意する複数のパラメータ（キーワード）の中から選択を行ったり、ユーザ I/F 8 を用いて概要の入力を行ったりすることで実現できる。

クライアント 1 が用意するパラメータとしては、例えば「青」、「赤」といった色情報や、明るさの情報、セグメントの面積比率、「丸」、「四角」などのセグメントの形状情報、画面の「上」、「右下」といった位置情報などが挙げられる。

このようなパラメータの組み合わせで、例えば「青」と「80%」を指定することにより、前述の「代表色が青で、かつフレーム画面中の 80% を占めるセグメント」の記述が入力されたこととなる。また、「赤

」色の「四角」い形状のセグメントが画面の「下」の方で「20%」を占めている、というように指定することにより、前述の「赤い車」の記述が入力されたこととなる。また、「赤い車」と「青い空」を含むビデオコンテンツ111等のように、複数のセグメントの特徴量を組み合わせて検索を行うこともできる。クライアント1が用意されたパラメータを選択した場合には、この選択結果が、そのまま検索キー108となり、検索処理部9から出力される。

第2図において、特徴復号部10は、検索処理部9からの検索指示107に従って、特徴ストリーム記憶部7から特徴ストリーム103を読み出して復号を行い、復号特徴量109を出力する。特徴ストリーム103は、第12図、第15図、第17図などのように、代表色125、サイズ127、セグメント数、キーフレーム番号122、縮小画像などの複数の特徴量が情報量を圧縮するために符号化されているため、同定のための特徴量の値を得るのに復号処理が必要となる。

特徴復号部10から出力された復号特徴量109は、特徴同定部11において検索キー108とのマッチング処理が行われる。第18図は実施の形態4による特徴同定部11の内部構成を示すブロック図である。図において、81a～81eは、検索キー108と復号特徴量109とのマッチング処理を行うマッチング処理部、82は、各マッチング処理部81a～81eの同定結果を総合的に判定し、マッチングの度合いの高いビデオコンテンツ111のポインタ110を出力するマッチング判定部である。

ここで、マッチング処理部81は複数用意され（81a～81e等）、それぞれ対応する特徴量についてマッチング処理を行う。例えば、マッチング処理部81aでは、復号された復号特徴量109の中から「青」の特徴量、また別のマッチング処理部81bでは、「80%」の特徴

量とのマッチングを行う。このとき、「水色」や「紺」といった「青」に近い色の特徴量や、「70%」や「90%」の特徴量に対しても、ユーザーが求めている映像に近い可能性があるため、特徴同定部11は、検索キー108とのマッチングを行う際に、完全に一致するものだけではなく、一致の度合いが高いものも候補とする。

各マッチング処理部81a～81eの同定結果は、マッチング判定部82に送られ、ここで各特徴量の一致の度合いが総合的に判定され、最終的にマッチング判定部82からの結果が、復号特徴量109と検索条件である検索キー108との一致度になる。一致の度合いを規定する閾値は、システムとして標準仕様のデフォルトの値が設定されるだけでなく、ユーザからも設定できるものとする（図示はしていない）。

特徴同定部11は、マッチングの度合いの高いビデオコンテンツ111のポインタ110を、サーバー2に対して出力し、サーバー2は、そのポインタ110に応じて、ビデオコンテンツ111をクライアント1に対して出力する。

クライアント1は、ビデオコンテンツ111をユーザI/F8に表示し、ユーザーが所望したビデオコンテンツ111であれば、検索処理を終了する。異なっている場合には、さらに別の検索キー108が生成されるように、ユーザーは再度パラメータの選択を行う。

ビデオコンテンツ111として、クライアント1に送られる画像データは、ビデオコンテンツ記憶部3に記憶されているビデオコンテンツ111の実態である必要はなく、前述の特徴ストリーム103の縮小画像（サムネイル画像）でも構わない。サムネイル画像を用いる場合には、サーバ2からクライアント1へ送られるビデオコンテンツ111のデータ量を削減でき、またユーザI/F8に表示できる画面のサイズは限定されているため、サムネイル画像であれば、複数の候補画像を同時に表

示することが可能であり、検索処理の操作性が向上する。

ビデオコンテンツ記憶部 3 に記憶されている画像が、あらかじめ限定されている場合には、ユーザ I / F 8 は、特徴ストリーム記憶部 7 に記憶されている特徴ストリーム 103 におけるサムネイル画像を、パラメータとして表示することも可能である。

以上のように、この実施の形態 4 によれば、上記のような検索処理を行うクライアント 1、特徴復号部 10、特徴同定部 11 により、ユーザは所望のビデオコンテンツ 111 の候補となるビデオコンテンツ 111 を、ほぼ自動的にかつ迅速に検索することが可能になる。特徴ストリーム 103 のデータ量は、一般に対応するビデオコンテンツ 111 よりも小さいので、特徴復号部 10 の処理も、ビデオ信号 102 を伸長復号する処理に比べれば極めて軽い処理である。

また、特徴ストリーム 103 の中に、サムネイル画像を含む場合は、多数の候補のビデオコンテンツ 111 の内容を一度に表示させ、かつブラウズするように構成することができるので検索の効率が著しく向上する。

なお、この実施の形態 4 では、第 2 図のシステムを前提とした検索処理を実施するクライアント 1 について説明したが、この実施の形態におけるクライアント 1 と同じ動作で、静止画像のライブラリから、所望の画像データを検索する処理を実現することも可能である。この場合は、特徴ストリーム 103 はビデオコンテンツ 111 ごとでなく、各静止画像ごとに作成される。

また、特徴ストリーム 103 は、ネットワークを通じて遠隔地に伝送することも可能であり、受信側で検索処理部 9 だけでなく、特徴符号化部 5 のような特徴ストリーム 103 の生成機能を備えていれば、既存の特徴ストリーム 103 の記述を書き換えて、新しい特徴ストリーム 10

3 を生成したり、それによって、ビデオコンテンツ 1 1 1 の表示規則を変更するなどのコンテンツ操作を行うことも可能である。もちろん、クライアント 1 とサーバ 2 の機能を併せ持つ装置を構成することも可能である。

実施の形態 5 .

第 2 図において、クライアント 1 が、上記の特徴同定部 1 1 において評価されたマッチングの度合いの情報を検索処理部 9 に入力して（図示せず）、表示する候補画像に対して、画像の順番や表示の大きさに重み付けを行うことにより、操作性を向上させることが可能である。第 1 9 図は優先度付き検索候補提示の例を示す図であり、特徴同定部 1 1 におけるマッチングの度合いに対応して、検索された候補コンテンツを、優先度をつけてユーザに提示している様子を示している。

以上のように、この実施の形態 5 によれば、優先度付き検索候補を提示することにより、ユーザは、より自分の検索要求に即したコンテンツを迅速に探し出すことができる。

実施の形態 6 .

この実施の形態では、ユーザ I / F 8 を介した他の検索条件入力方法について説明する。ユーザ I / F 8 を用いて、概要の入力を行う方法としては、マウス等のポインターデバイスを用いて、直接形状を書き込んだり、色を塗ったりして、おおまかなスケッチを描くような方法が考えられる。

例えば、第 1 9 図の第 2 候補のセグメントに示すように、画面の下の方に大きく「赤」色を描いたり、第 3 候補のセグメントに示すように、画面の左の方に小さく「赤い車」の絵を描いたり、第 4 候補のセグメン

トに示すように、画面の右のように、大きく「青」色を描いたりする。
第2図において、ユーザ指示・設定105により、ユーザの直感的な概要入力を行った場合には、検索キー108を抽出するための処理が、検索処理部9において行われる。

検索処理部9では、概要入力から、色情報などを用いて個々のセグメントに分割し、その色が描かれた面積を算出したり、画面内の位置を求めたりする。このような処理により、「青」、「赤」といった色情報や、各色情報の面積比率、各色の形状や位置などが抽出されて、検索キー108として出力される。

以上のように、この実施の形態6によれば、ユーザの直感的な概要入力を行うことにより、ビデオコンテンツ111を効率的に検索することができる。

実施の形態7.

実施の形態3で述べたセグメントの動き情報171が抽出されている場合には、検索キー108として、動き情報171を用いることが可能である。ユーザI/F8は、選択できるパラメータとして、「左から右へ」、「上から下へ」、「ズームイン」といった動き情報171を用意しておく。また、映像信号の時間的変化が抽出されている場合には、色の変化や明るさの変化のパラメータを用意しておき、ユーザーが選択できるようにする。

ユーザーが映像内容の概要入力を行う際に、1つの入力のみを行うのではなく、例えば、入力を2回行い、その2つの画像間の時間を入力できれば、検索処理部9は、2つの入力画像とその時間間隔から、オブジェクトの動き量や、映像信号の時間変化情報を抽出して、検索キー108を生成することができる。

以上のように、この実施の形態 7 によれば、動き情報 1 7 1 を用いてユーザが求めるビデオコンテンツ 1 1 1 を検索することができる。

産業上の利用可能性

以上のように、この発明に係る画像特徴符号化方法及び画像検索方法は、多くのユーザが様々な検索ツールを用いて、画像を検索するのに適している。

請 求 の 範 囲

1. 画像フレームから所定の画像領域であるセグメントを抽出し、抽出した上記セグメントごとに上記セグメントを識別するセグメント番号を付与し、上記セグメントに代表色を割り当て、上記セグメントの上記画像フレームに対する面積率を算出し、上記セグメントごとに、上記代表色及び上記面積率を上記セグメントの特徴量として、上記セグメント番号と共に符号化して特徴ストリームを生成することを特徴とする画像特徴符号化方法。

2. 画像フレームからセグメントを抽出する際に、色情報に基づく領域分類により上記セグメントを抽出し、上記セグメントに代表色を割り当てる際に、上記セグメントの抽出に用いた上記色情報を上記代表色として割り当てることを特徴とする請求の範囲第1項記載の画像特徴符号化方法。

3. 画像フレームからセグメントを抽出する際に、隣接する画像フレーム間で上記セグメントの同定を行い、同じセグメントと判定されたセグメントに対して同一のセグメント番号を付与することを特徴とする請求の範囲第1記載の画像特徴符号化方法。

4. 画像フレームからセグメントを抽出する際に、隣接する画像フレーム間でセグメントの追跡を行い、同じセグメントと判定されたセグメントに対して動き情報を求め、上記セグメントごとに、上記動き情報を上記セグメントの特徴量として符号化して特徴ストリームを生成することを特徴とする請求の範囲第1記載の画像特徴符号化方法。

5. ビデオ信号のフレーム群から検索の手がかりとして適切な画像フレームをキーフレームとして抽出し、抽出したキーフレームからセグメントを抽出することを特徴とする請求の範囲第1項記載の画像特徴符号化方法。

6. キーフレームの所定領域ごとの画素値の平均値を求めて上記キーフレームの縮小画像を生成し、上記キーフレームごとに、上記縮小画像を上記キーフレームの特徴量として符号化して特徴ストリームを生成することを特徴とする請求の範囲第1記載の画像特徴符号化方法。

7. 画像フレームを記憶する第1の記憶部と、上記第1の記憶部に記憶している画像フレームの特徴量を符号化して特徴ストリームとして記憶する第2の記憶部を備えて画像検索を行う画像検索方法において、ユーザからの検索指示に基づき、上記第2の記憶部に記憶している上記特徴量を復号し、ユーザからの検索条件と上記復号した特徴量との同定を行い、上記第1の記憶部に記憶している画像フレームを検索することを特徴とする画像検索方法。

8. 第2の記憶部に記憶している特徴量として、画像フレームにおける所定の画像領域であるセグメントの代表色を含み、ユーザからの検索条件として上記代表色を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

9. 第2の記憶部に記憶している特徴量として、画像フレームにおける所定の画像領域であるセグメントの上記画像フレームに対する面積率を

含み、ユーザからの検索条件として上記面積率を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

10. 第2の記憶部に記憶している特徴量として、隣接する画像フレームの動き情報を含み、ユーザからの検索条件として上記動き情報を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

11. 第2の記憶部に記憶している特徴量として、画像フレームの縮小画像を含み、ユーザからの検索条件と上記復号した特徴量との同定を行い、上記縮小画像を上記ユーザに提示することを特徴とする請求の範囲第7項記載の画像検索方法。

12. 第2の記憶部に記憶している特徴量として、指定された被写体が画像フレームに存在するか否かの情報を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

13. 第2の記憶部に記憶している特徴量として、指定された被写体が以後の画像フレームに存在するか否かの情報を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

14. 第2の記憶部に記憶している特徴量として、指定された被写体が以前の画像フレームに存在するか否かの情報を含むことを特徴とする請求の範囲第7項記載の画像検索方法。

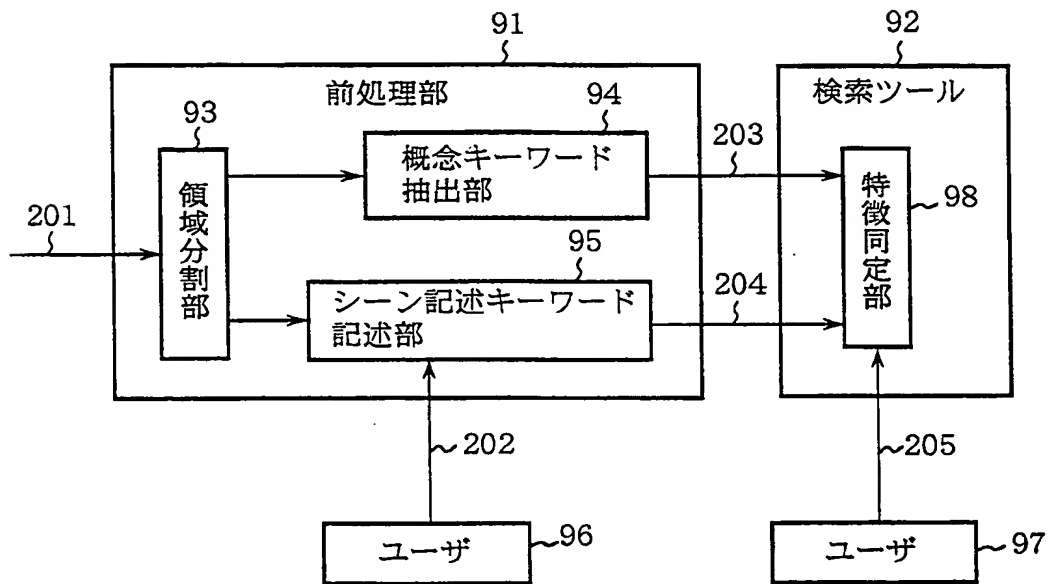
15. ユーザからの検索条件と復号した特徴量との同定を行う際に、優先度をユーザに提示することを特徴とする請求の範囲第7項記載の画像

検索方法。

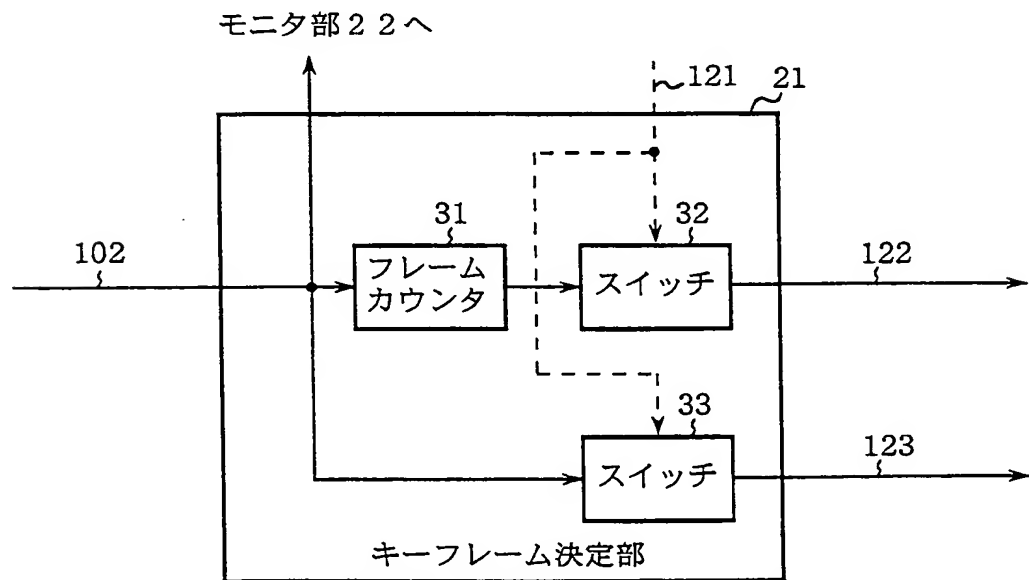
16. ユーザからの検索条件と復号した特徴量との同定を行う際に、ユーザからの複数の検索条件と復号した複数の特徴量との同定を行い、総合的に判断して、第1の記憶部に記憶している画像フレームを検索することを特徴とする請求の範囲第7項記載の画像検索方法。

1/14

第 1 図

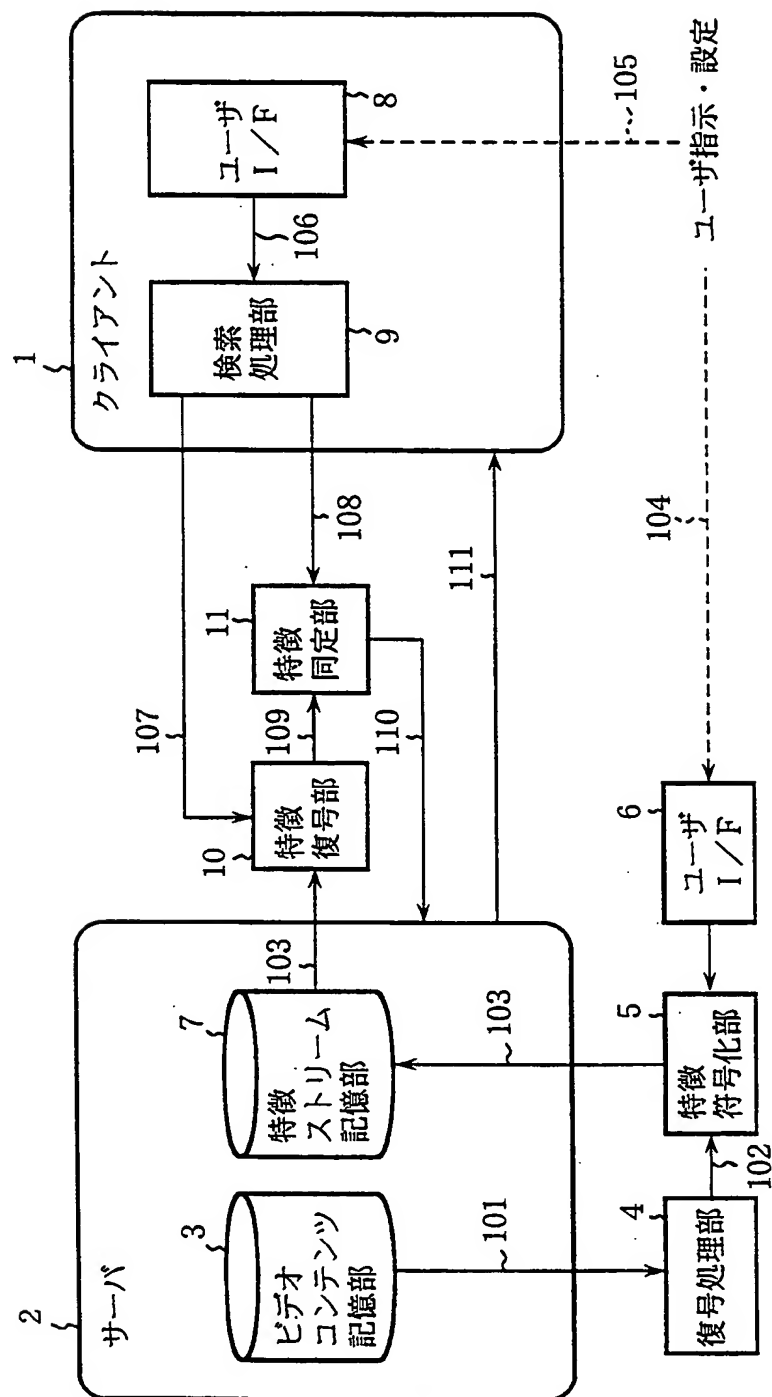


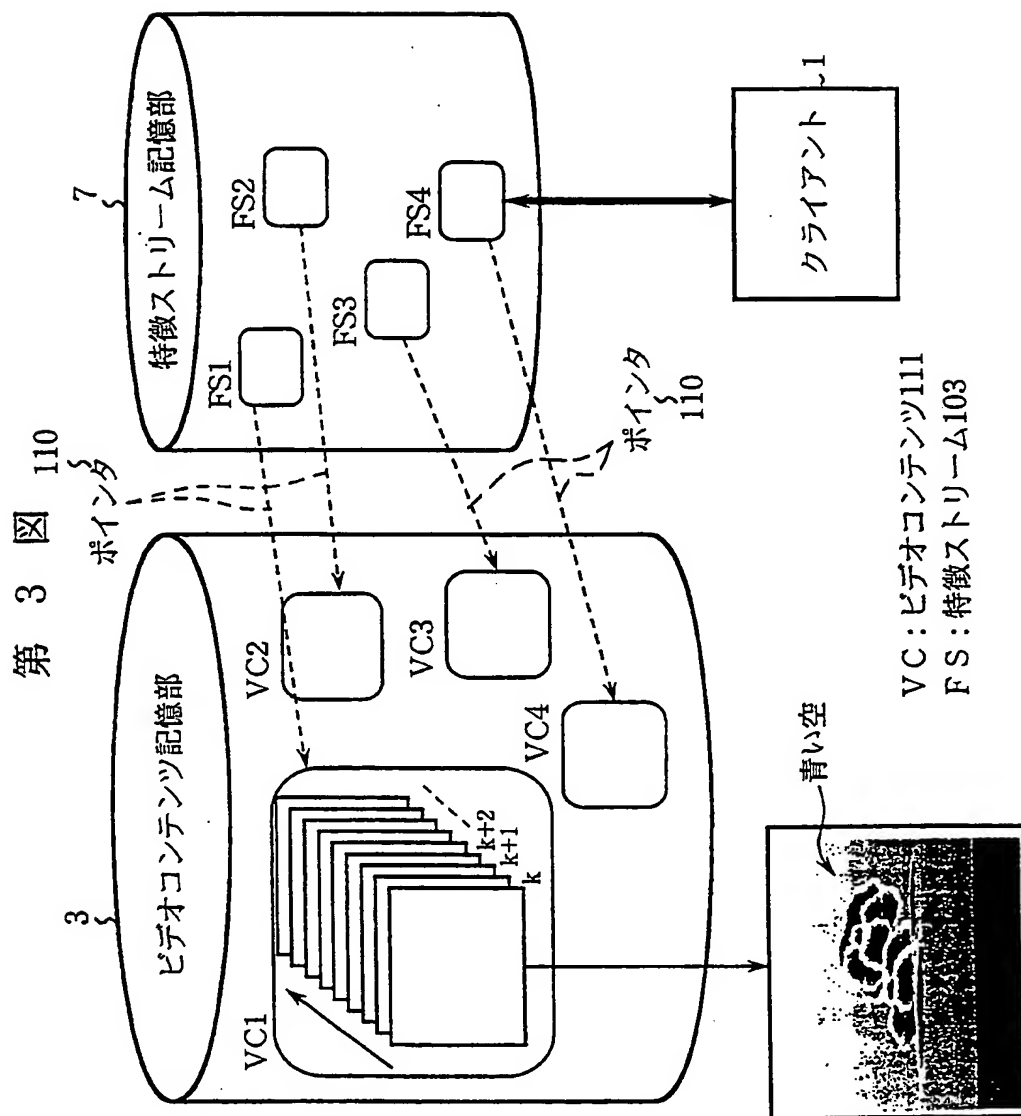
第 6 図



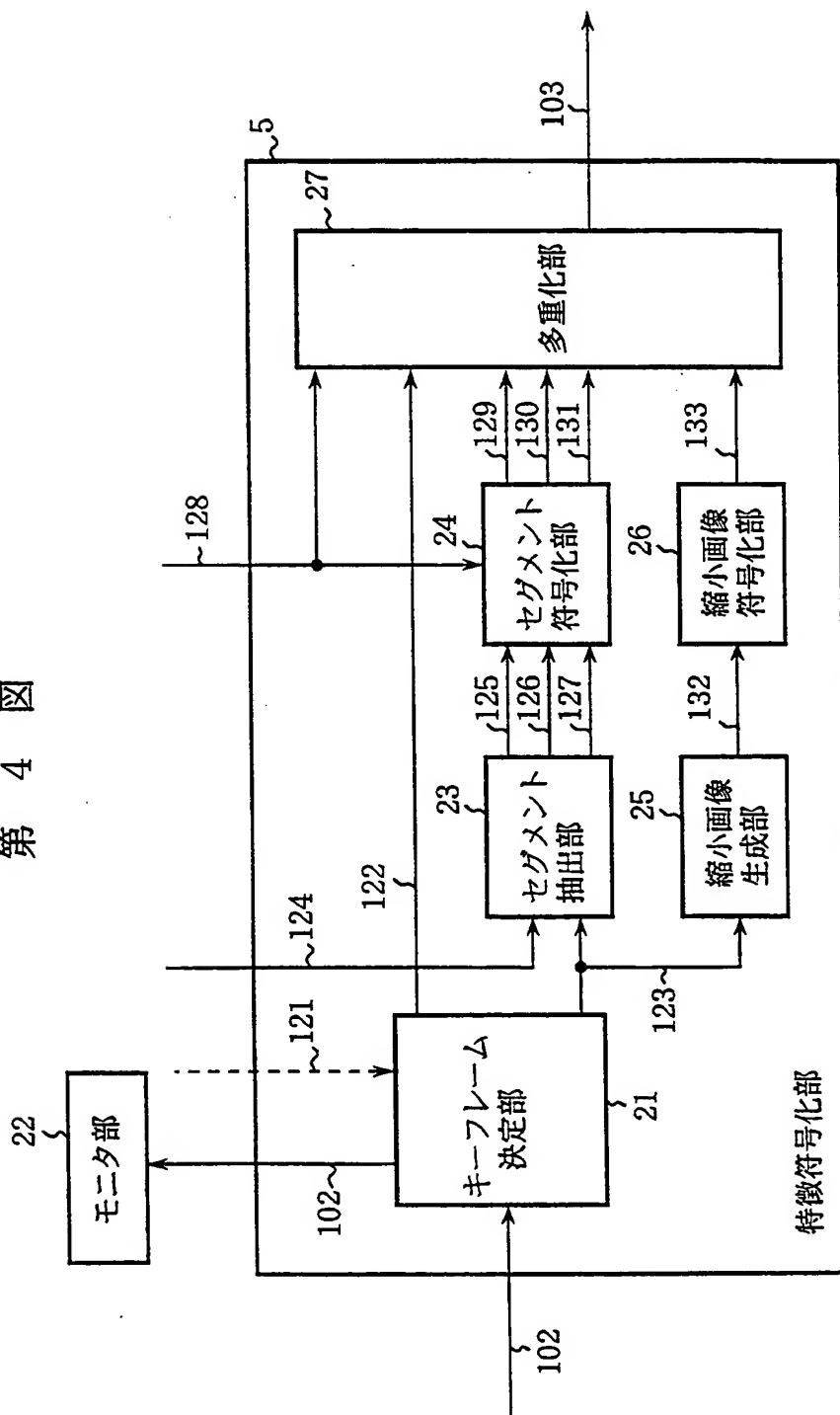
2/14

第 2 図



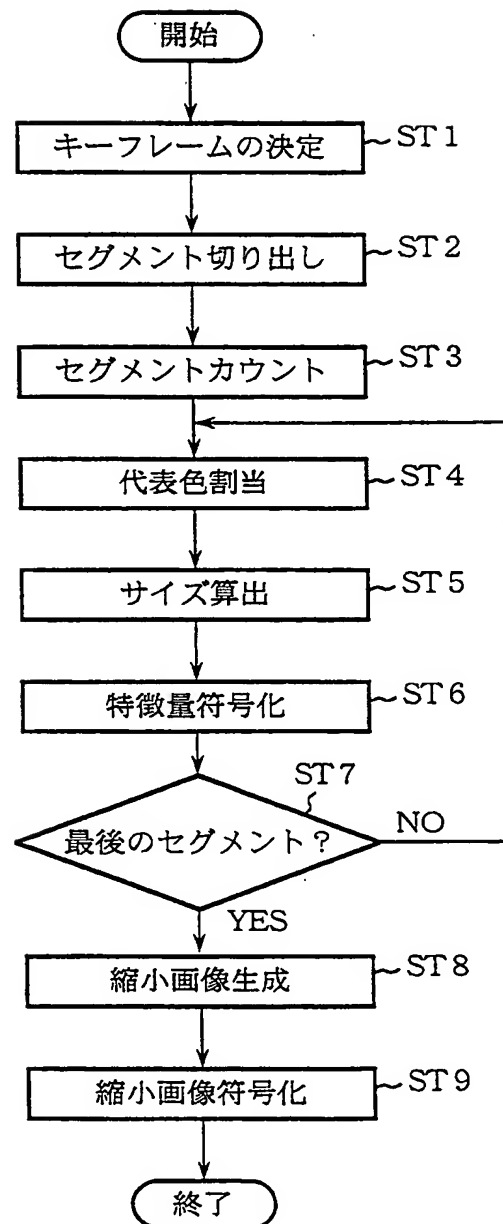


第 4 圖



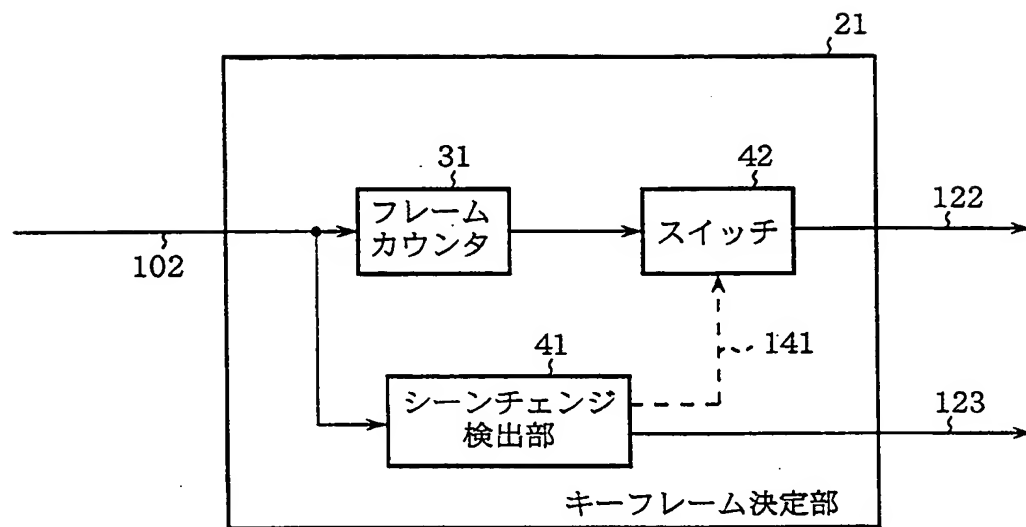
5/14

第 5 図

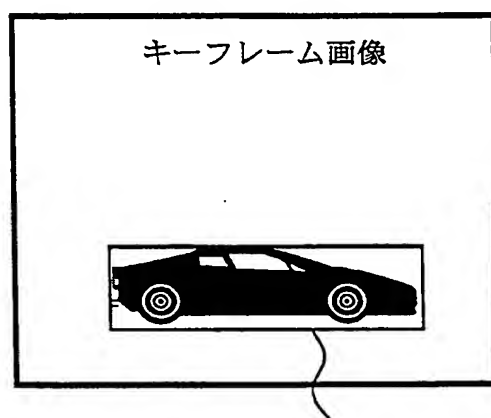


6/14

第 7 図

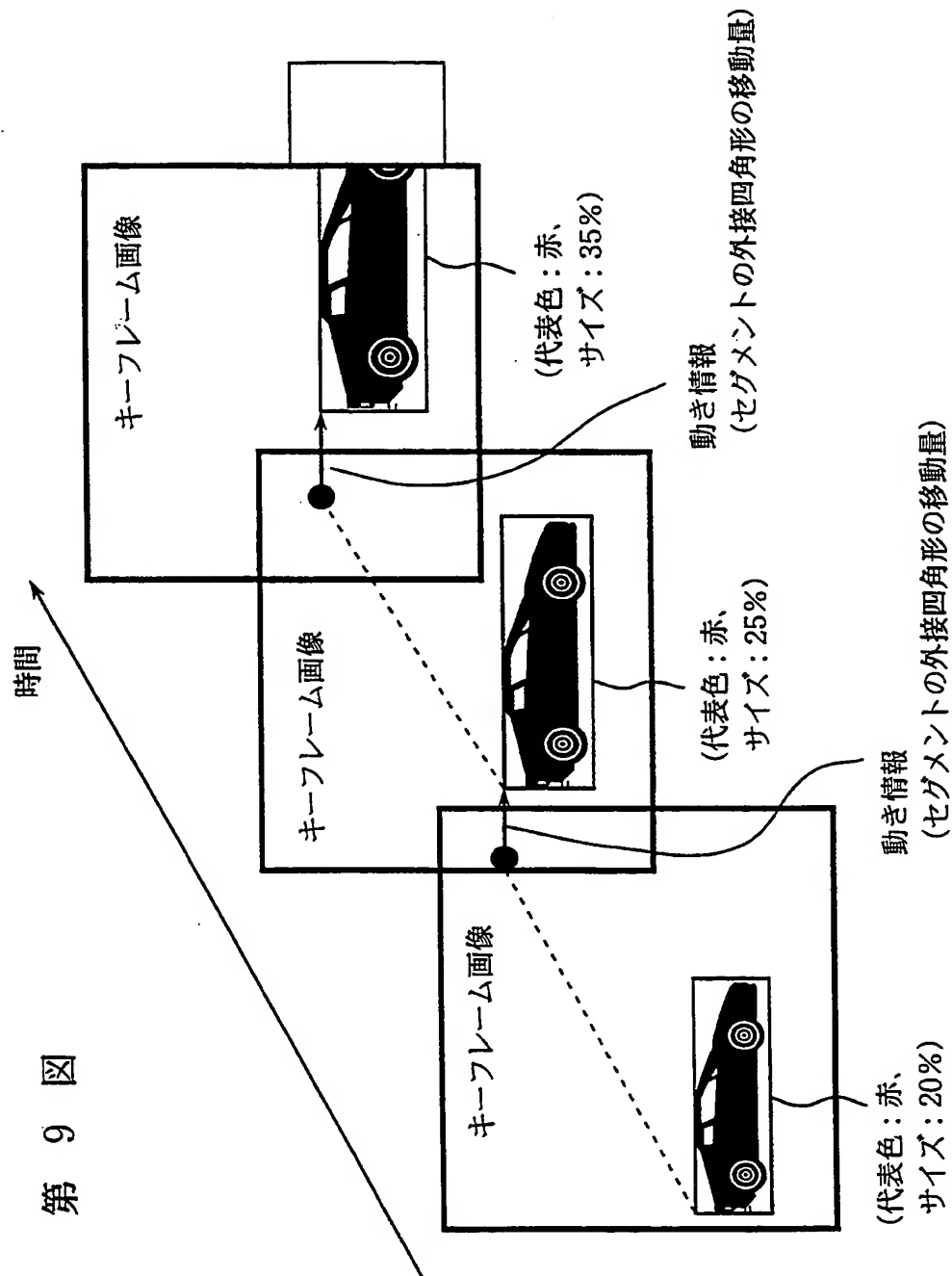


第 8 図

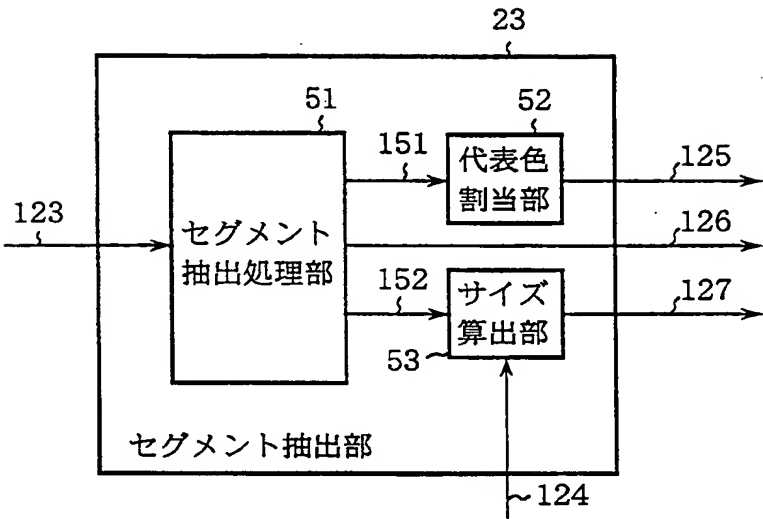


セグメント(代表色：赤，サイズ：25%)

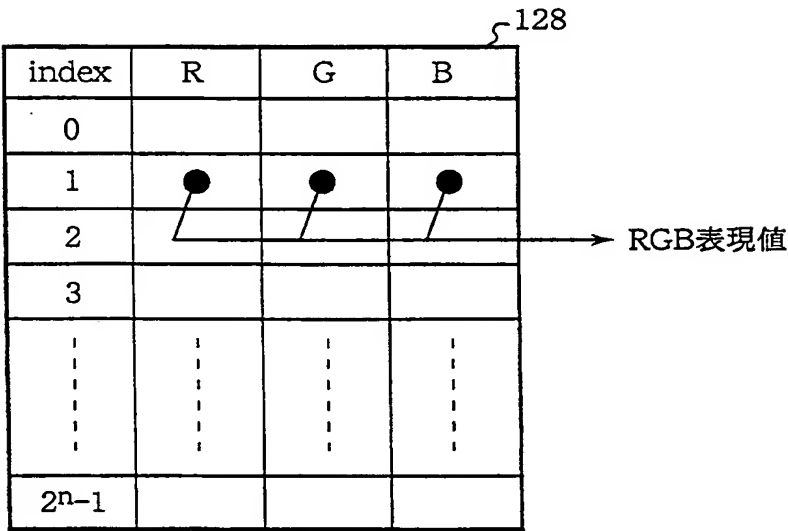
7/14



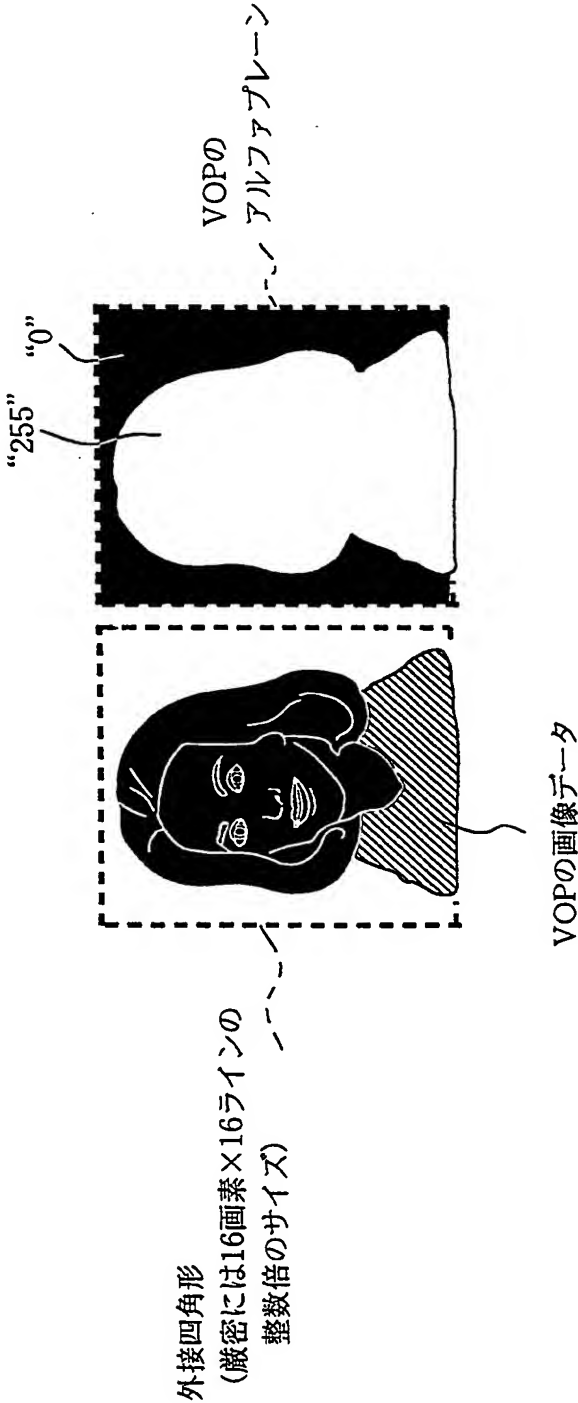
第 10 図



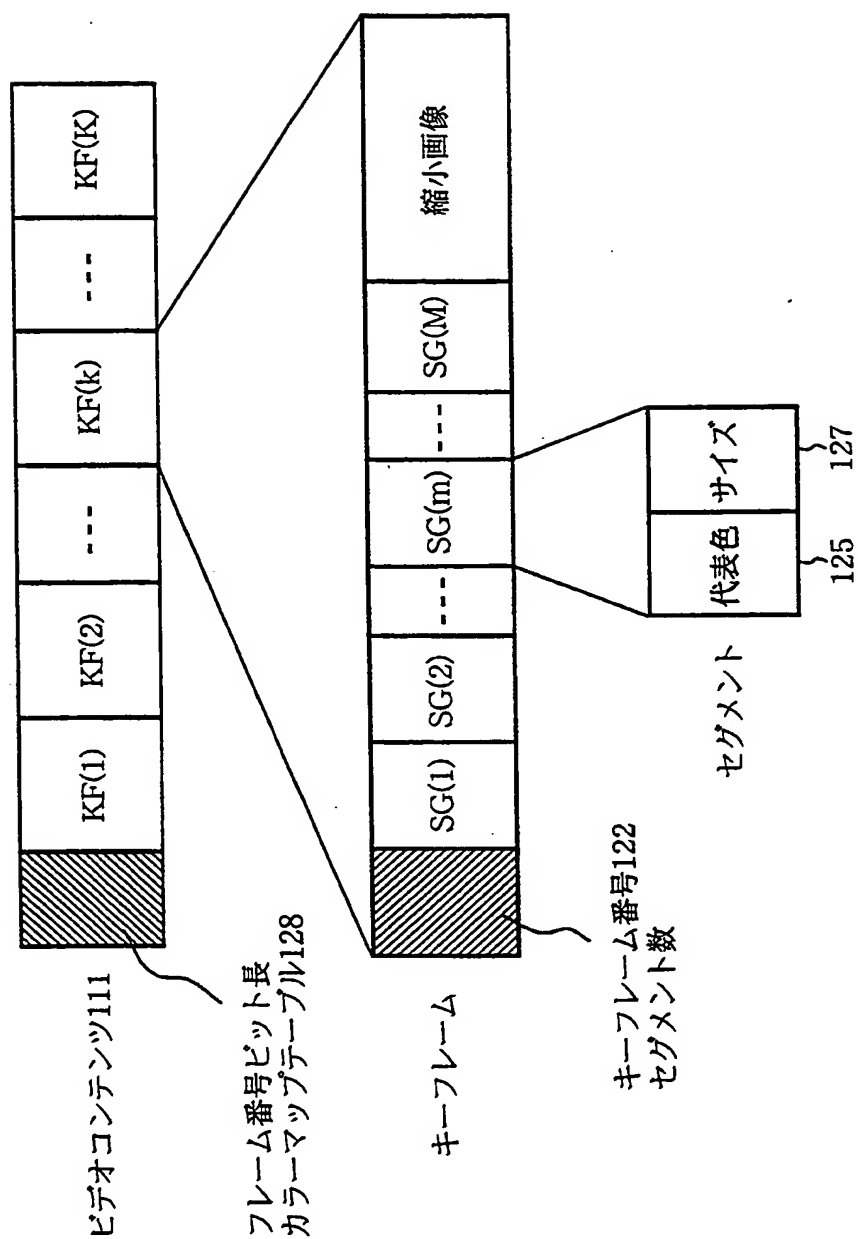
第 13 図



第 11 図

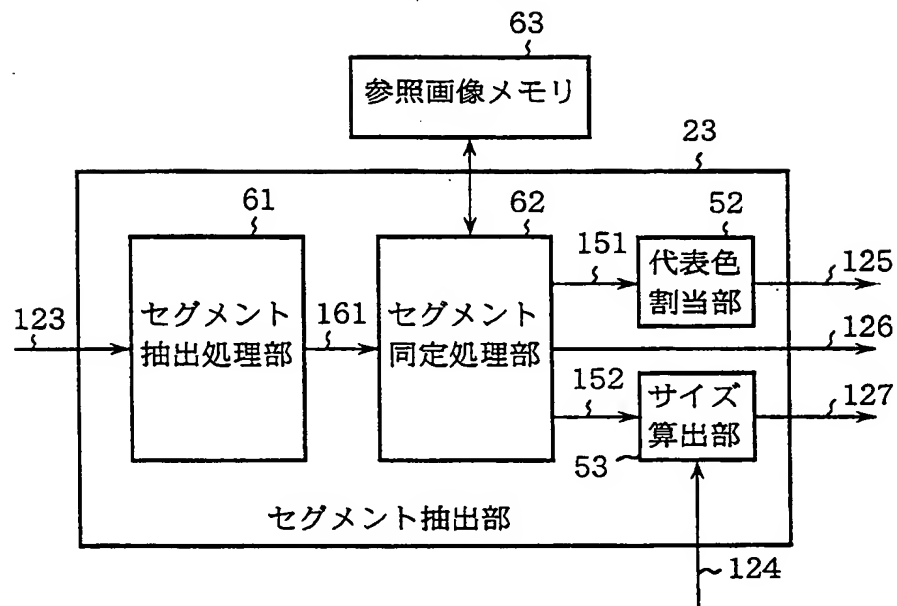


第 12 図

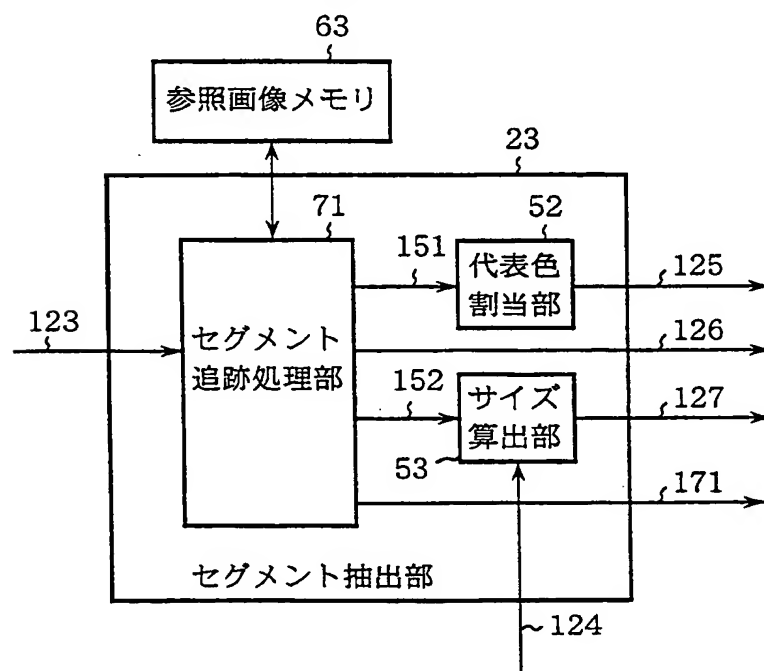


11/14

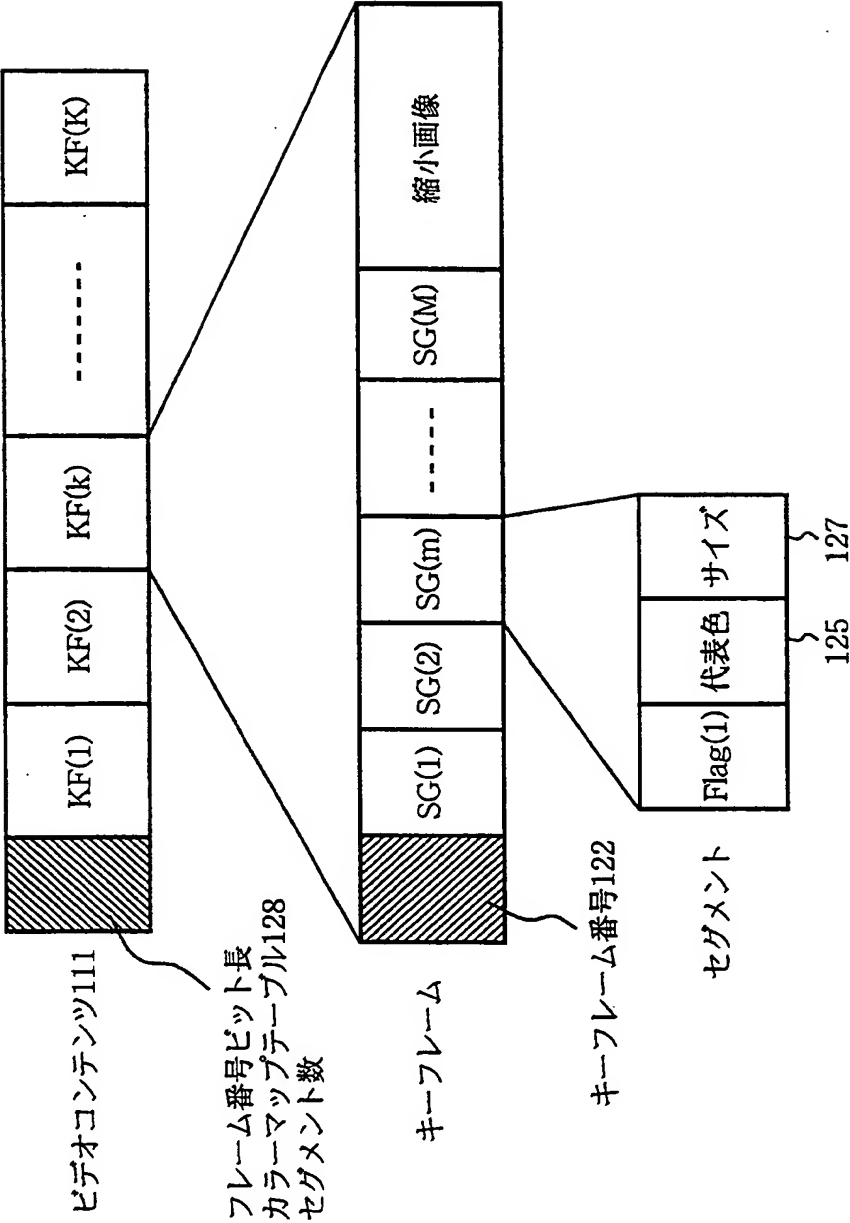
第 14 図



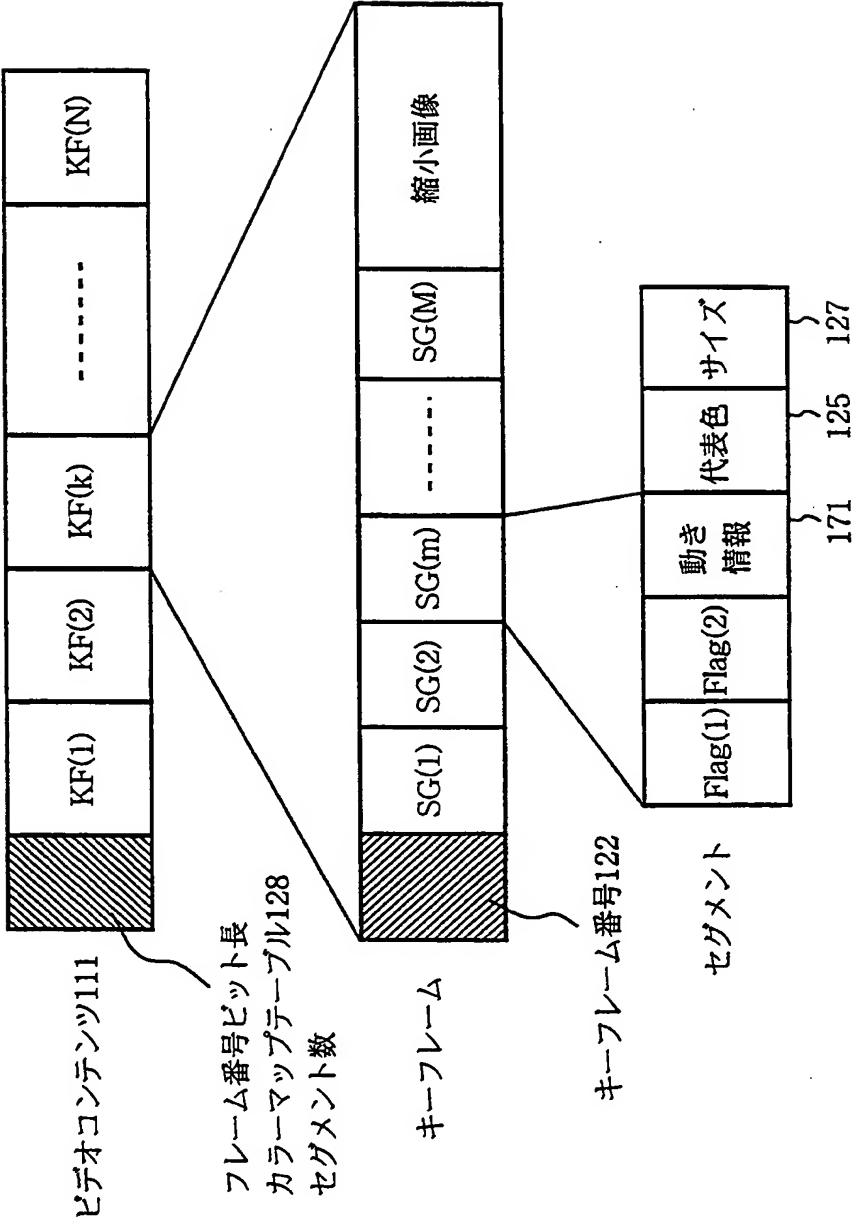
第 16 図



第 15 図

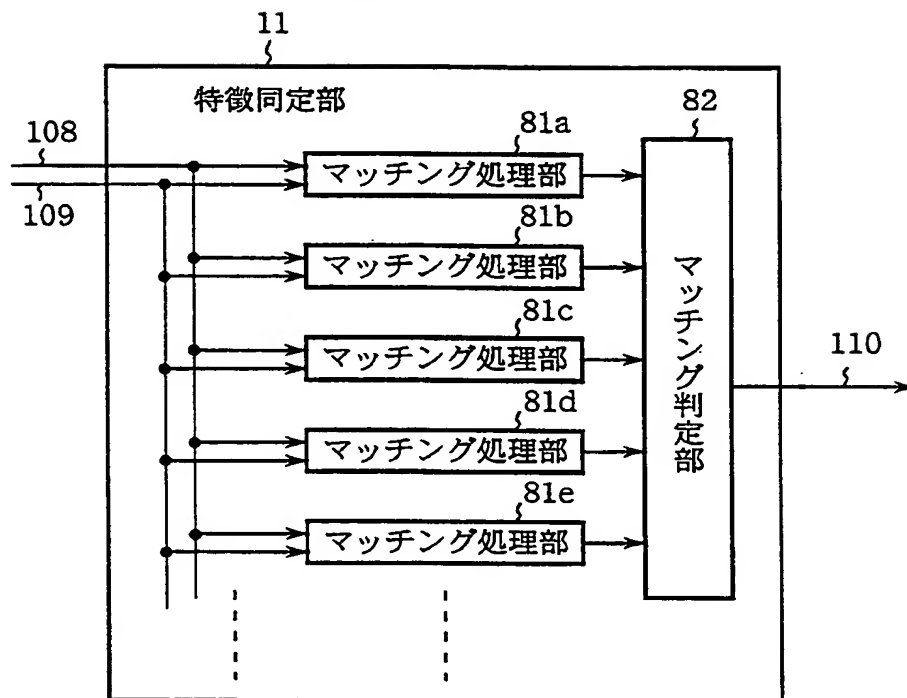


第 17 図

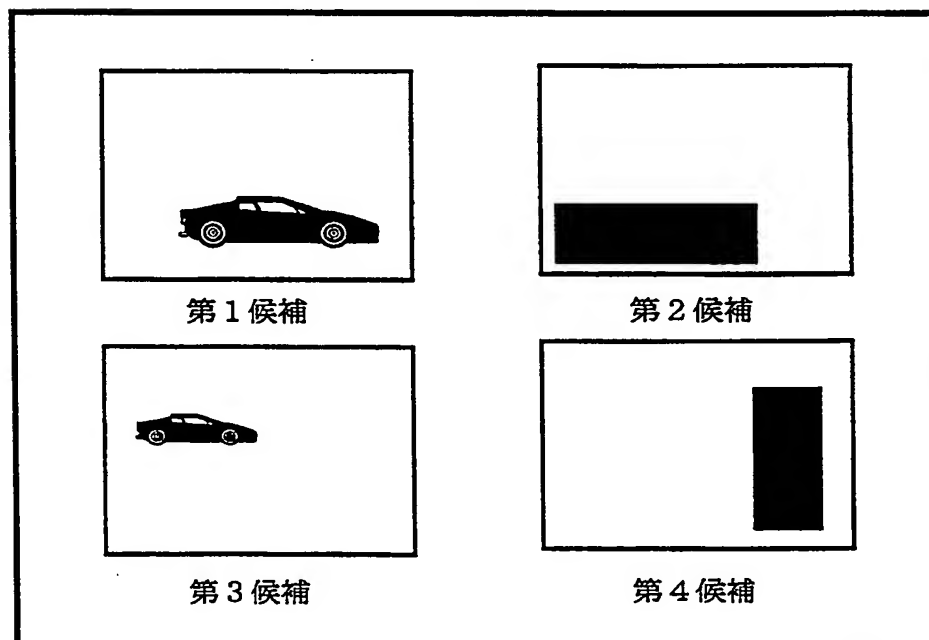


14/14

第 18 図



第 19 図



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP99/00403

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl.⁶ H04N7/24, 5/93

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl.⁶ H04N7/24-7/68, H04N5/91-5/956

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1926-1999
Kokai Jitsuyo Shinan Koho 1971-1999

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Tanaka et al., "MPEG stream o jiyuujizai ni ayatsuru", Nikkei Electronics, No. 711 9 March, 1998 (Tokyo) pp.149-154	1-16
Y	Ono et al., "Joutai seni model to scene kijutsu gengo ni yoru jidou key word fuyo kinou o motsu gazou data base no hyouka", Denshi Jouhou Tsuushin Gakkai Rombunshi D-II Vol. J79-D-II No. 4, April, 1996 (Tokyo) pp.476-483	1-16
A	Norikazu Miki, "Tasai na eizou, onsei o jizai ni fugouka suru MPEG-4 no subete", Kogyo Chosakai 30 September, 1998 (30. 09. 98) pp.283-290 5.13.4 "Application of MPEG-7"	1-16
A	Nagasaka et al., "Cut no jikeiretsu coding ni motozuku eizou scene no jitsujikan shikibetsuhou", Denshi Jouhou Tsuushin Gakkai Rombunshi D-II Vol. J79-D-II No. 4, April, 1996 (Tokyo) pp.531-537	1-16

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>
--	---

Date of the actual completion of the international search
21 April, 1999 (21. 04. 99)

Date of mailing of the international search report
11 May, 1999 (11. 05. 99)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP99/00403**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	A.A. Alatan et al., "Image Sequence Analysis for Emerging Interactive Multimedia Services-European COST 211 Framework" IEEE TRANSACTION OF CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL.8 NO.7, November, 1998 pp.802-813	1-16

国際調査報告

国際出願番号 PCT/J P 99/00403

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. cl⁸ H04N 7/24, 5/93

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. cl⁸ H04N 7/24-7/68Int. cl⁸ H04N 5/91-5/956

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1926-1999年

日本国公開実用新案公報 1971-1999年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

INSPEC

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	田中他、「MPEGストリームを自由自在に操る」、日経エレクトロニクス、No. 711 9.3月.1998 (東京) pp. 149-154	1-16
Y	小野他 「状態遷移モデルとシーン記述言語による自動キーワード付与機能をもつ画像データベースの評価」、電子情報通信学会論文誌 D-II Vol. J79-D-II No. 4, 4月.1996 (東京) pp. 476-483	1-16
A	三木弼一「多彩な映像、音声を自在に符号化するMPEG-4のすべて」工業調査会 30.9月.1998(30.09.98) pp. 283-290 (5.13.4 MPEG-7の応用)	1-16

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

21.04.99

国際調査報告の発送日

11.05.99

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

松永 隆志

5 P

4 2 2 8

電話番号 03-3581-1101 内線 6979

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	長坂他 「カットの時系列コーディングに基づく映像シーンの実時間識別法」 電子情報通信学会論文誌 D-II Vol. J79-D-II No. 4, 4月. 1996 (東京) pp. 531-537	1-16
A	A. A. Alatan et al 「Image Sequence Analysis for Emerging Interactive Multimedia Services-European COST 211 Framework」 IEEE TRANSACTION ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 8 NO. 7, 11月. 1998 pp. 802-813	1-16